

Christine Erbe
Jeanette A. Thomas *Editors*

# Exploring Animal Behavior Through Sound: Volume 1

## Methods

OPEN ACCESS

ASA PRESS

Springer

# Exploring Animal Behavior Through Sound: Volume 1

Christine Erbe • Jeanette A. Thomas
Editors

# Exploring Animal Behavior Through Sound: Volume 1

## Methods

Springer

ASA PRESS

*Editors*
Christine Erbe
Centre for Marine Science
and Technology
Curtin University
Perth, WA, Australia

Jeanette A. Thomas (*deceased*)
Moline, IL, USA



ISBN 978-3-030-97538-8 ISBN 978-3-030-97540-1 (eBook)
https://doi.org/10.1007/978-3-030-97540-1

© Springer Nature Switzerland AG 2022. This book is an open access publication.
Jointly published with ASA Press
**Open Access** This book is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.
The images or other third party material in this book are included in the book's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the book's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publishers, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publishers nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publishers remain neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Cover photo: Acoustic recording of an Adélie penguin colony at Brown Bluff, Antarctic Sound (© Ole Næsbye Larsen)

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

*In loving memory of*
*Jeanette A. Thomas,*
*A pioneer of animal bioacoustics,*
*A role model, mentor, colleague,*
*And dear friend to many of us.*
*We miss you, Jeanette.*

**The ASA Press**

ASA Press, which represents a collaboration between the Acoustical Society of America and Springer Nature, is dedicated to encouraging the publication of important new books as well as the distribution of classic titles in acoustics. These titles, published under a dual ASA Press/Springer imprint, are intended to reflect the full range of research in acoustics. ASA Press titles can include all types of books that Springer publishes, and may appear in any appropriate Springer book series.

*Editorial Board*

Mark F. Hamilton (Chair), University of Texas at Austin
James Cottingham, Coe College
Timothy F. Duda, Woods Hole Oceanographic Institution
Robin Glosemeyer Petrone, Threshold Acoustics
William M. Hartmann (Ex Officio), Michigan State University
Darlene R. Ketten, Boston University
James F. Lynch (Ex Officio), Woods Hole Oceanographic Institution
Philip L. Marston, Washington State University
Arthur N. Popper (Ex Officio), University of Maryland
Christine H. Shadle, Haskins Laboratories
G. Christopher Stecker, Boys Town National Research Hospital
Stephen C. Thompson, The Pennsylvania State University
Ning Xiang, Rensselaer Polytechnic Institute



# Preface

The idea for this textbook on *Animal Bioacoustics* was Jeanette's. She reached out to bioacousticians working on the different animal taxa and received great interest in this book. Experts from around the globe joined her effort, developing chapters on bioacoustic studies on the diverse animal taxa, from invertebrates and insects, to amphibians, reptiles, fishes, birds, and mammals. It soon became obvious that the developing chapters relied on common background knowledge, techniques, and terminology. The need for a volume on methods to precede the volume on taxon-specific bioacoustic studies was identified and this is when I came onboard.

In this volume, Chapter 1 presents a brief history to bioacoustic recording and equipment. Chapter 2 provides guidance on choosing and calibrating equipment. Chapter 3 explains how to collect bioacoustic data in the field and laboratory, and what metadata are important to document. Chapter 4 introduces basic acoustic concepts, standard terminology, quantities and units, and basic signal processing methods. Chapter 5 delves into the source–path–receiver model, applied to terrestrial bioacoustic studies, with a comprehensive treatise of sound propagation in terrestrial environments. Chapter 6 is devoted to the intricacies of sound propagation under water. Chapter 7 explores terrestrial and aquatic soundscapes and introduces basic analysis tools. Chapter 8 gives an overview of software algorithms for automated detection and classification of animal sounds. Chapter 9 unravels analytical and statistical methods for analyzing bioacoustic data. Chapter 10 presents behavioral and physiological methods for studying animal hearing. The final three chapters apply the tools presented in the first ten chapters to taxon-overarching topics. Chapter 11 explores animal acoustic and vibrational communication. Chapter 12 provides an overview of echolocation in bats, dolphins, birds, and shrews. And Chap. 13 gives examples of the effects of noise on animals.

The intended audience includes students and researchers of animal ecology and, specifically, animal behavior, who wish to add acoustics to their toolbox. Environmental managers in industry and government, members of non-governmental organizations concerned with animal conservation, and regulators of noise might equally find the book useful. The book will empower its readers to understand and apply the bioacoustic research literature, design their own studies in the field and laboratory, avoid common pitfalls and mistakes, choose appropriate equipment, apply different data

analysis methods, correctly interpret their data, adequately archive data for future applications, and apply their results to management and conservation.

I would like to thank Keith Attenborough, Jay Barlow, Ross Chapman, Russ Charif, Kurt Fristrup, Karl-Heinz Frommolt, Bob Gisiner, Alan Grinnell, Shane Guan, Shizuko Hiryu, Dorian Houser, Vincent Janik, Colleen LePrell, Peter Narins, Eric Rexstad, James Simmons, Hans Slabbekoorn, and Meta Virant-Doberlet for reviewing one or more chapters in this volume.

A special thank-you goes to Lars Koerner at Springer Verlag in Heidelberg for his emotional, technical, and editorial support throughout the years, in particular the final year.

Open access to this book was mostly funded by the Richard Lounsbery Foundation, as a contribution to the International Quiet Ocean Experiment. The remainder of fees was covered by the Centre for Marine Science and Technology at Curtin University, the Cornell Lab of Ornithology, and l'Université de Toulon. Thank you!

Jeanette A. Thomas was a pioneer of animal bioacoustics. She successfully straddled both terrestrial and aquatic worlds, studying animals from the tropics to the poles. This book is a testament to her legacy.

Perth, WA
September 2021

Christine Erbe

The original online version of this Frontmatter and backmatter was revised.

# Contents


**1 History of Sound Recording and Analysis Equipment** . . . . . 1
Gianni Pavan, Gregory Budney, Holger Klinck, Hervé Glotin, Dena J. Clink, and Jeanette A. Thomas
1.1 Introduction . . . . . 1
1.2 Advances in Recorders . . . . . 2
1.3 Advances in Microphones . . . . . 14
1.4 Advances in Hydrophones . . . . . 20
1.5 Autonomous Mobile Systems . . . . . 24
1.6 Advances in Sound Analysis Hard- and Software . . . . . 27
1.7 Summary . . . . . 32
References . . . . . 33

**2 Choosing Equipment for Animal Bioacoustic Research** . . . . . 37
Shyam Madhusudhana, Gianni Pavan, Lee A. Miller, William L. Gannon, Anthony Hawkins, Christine Erbe, Jennifer A. Hamel, and Jeanette A. Thomas
2.1 Introduction . . . . . 37
2.2 Basic Concepts of Sound Recording . . . . . 38
2.3 Instrumentation of Signal Chain Components . . . . . 40
2.4 Autonomous Recorders . . . . . 63
2.5 Recording Directly to a Computer . . . . . 66
2.6 Calibration . . . . . 66
2.7 Other Gear . . . . . 73
2.8 Summary . . . . . 80
2.9 Additional Resources . . . . . 81
References . . . . . 81

**3 Collecting, Documenting, and Archiving Bioacoustical Data and Metadata** . . . . . 87
William L. Gannon, Rebecca Dunlop, Anthony Hawkins, and Jeanette A. Thomas
3.1 Introduction . . . . . 87
3.2 Ethical Research . . . . . 88
3.3 Good Practices in Bioacoustical Studies . . . . . 89
3.4 Playback Methods and Controls . . . . . 94
3.5 Considerations for Terrestrial Field Studies . . . . . 99
3.6 Considerations for Aquatic Field Studies . . . . . 99

3.7 Considerations for Studies on Captive Animals . . . . . 102
3.8 Digital File Format . . . . . 104
3.9 Data Storage . . . . . 105
3.10 Archiving Recordings . . . . . 105
3.11 Repositories of Bioacoustical Data . . . . . 106
3.12 Summary . . . . . 107
3.13 Additional Resources . . . . . 108
References . . . . . 108

**4 Introduction to Acoustic Terminology and Signal Processing** . . . . . 111
Christine Erbe, Alec Duncan, Lauren Hawkins, John M. Terhune, and Jeanette A. Thomas
4.1 What Is Sound? . . . . . 111
4.2 Terms and Definitions . . . . . 112
4.3 Acoustic Signal Processing . . . . . 134
4.4 Localization and Tracking . . . . . 142
4.5 Symbols and Abbreviations (Table 4.10) . . . . . 149
4.6 Summary . . . . . 150
References . . . . . 150

**5 Source-Path-Receiver Model for Airborne Sounds** . . . . . 153
Ole Næsbye Larsen, William L. Gannon, Christine Erbe, Gianni Pavan, and Jeanette A. Thomas
5.1 Introduction . . . . . 153
5.2 Sound Propagation in Terrestrial Environments . . . . . 155
5.3 The Source-Path-Receiver Model for Animal Acoustic Communication . . . . . 171
5.4 Summary . . . . . 178
5.5 Additional Resources . . . . . 179
References . . . . . 179

**6 Introduction to Sound Propagation Under Water** . . . . . 185
Christine Erbe, Alec Duncan, and Kathleen J. Vigness-Raposa
6.1 Introduction . . . . . 185
6.2 The Sonar Equation . . . . . 186
6.3 The Layered Ocean . . . . . 190
6.4 Propagation Loss . . . . . 193
6.5 Practical Acoustic Modeling Examples . . . . . 208
6.6 Summary . . . . . 214
6.7 Additional Resources . . . . . 214
References . . . . . 214

**7 Analysis of Soundscapes as an Ecological Tool** . . . . . 217
Renée P. Schoeman, Christine Erbe, Gianni Pavan, Roberta Righini, and Jeanette A. Thomas
7.1 Introduction . . . . . 217
7.2 Terrestrial Soundscapes . . . . . 219

7.3 Aquatic Soundscapes . . . 227
7.4 Soundscape Changes Over Space and Time . . . 235
7.5 How to Analyze Soundscapes . . . 239
7.6 Applications of Soundscape Studies . . . 249
7.7 Conclusion . . . 251
7.8 Additional Resources . . . 252
References . . . 254

**8 Detection and Classification Methods for Animal Sounds** . . . 269
Julie N. Oswald, Christine Erbe, William L. Gannon, Shyam Madhusudhana, and Jeanette A. Thomas
8.1 Introduction . . . 269
8.2 Qualitative Naming and Classification of Animal Sounds . . . 271
8.3 Detection of Animal Sounds . . . 276
8.4 Quantitative Classification of Animal Sounds . . . 282
8.5 Challenges in Classifying Animal Sounds . . . 299
8.6 Summary . . . 304
8.7 Additional Resources . . . 306
References . . . 307

**9 Fundamental Data Analysis Tools and Concepts for Bioacoustical Research** . . . 319
Chandra Salgado Kent, Tiago A. Marques, and Danielle Harris
9.1 Introduction . . . 319
9.2 Developing a Clear Research Question . . . 321
9.3 Designing the Study and Collecting Data . . . 321
9.4 Data Types and Statistical Concepts . . . 325
9.5 Tackling Analyses . . . 331
9.6 Examples in Bioacoustics . . . 347
9.7 Software for Analyses . . . 351
9.8 Summary . . . 351
References . . . 352

**10 Behavioral and Physiological Audiometric Methods for Animals** . . . 355
Sandra L. McFadden, Andrea Megela Simmons, Christine Erbe, and Jeanette A. Thomas
10.1 Introduction . . . 355
10.2 What Is an Audiogram? . . . 356
10.3 Behavioral Methods for Audiometric Studies on Live Animals . . . 359
10.4 Physiological Methods for Audiometric Studies on Live Animals . . . 372
10.5 Other Audiometric Measurements . . . 379
10.6 Summary . . . 383
References . . . 384

**11 Vibrational and Acoustic Communication in Animals** . . . . . 389
Rebecca Dunlop, William L. Gannon, Marthe Kiley-Worthington, Peggy S. M. Hill, Andreas Wessel, and Jeanette A. Thomas
11.1 Introduction . . . . . 389
11.2 The Origins of Substrate-Borne Vibrational and Acoustic Communication . . . . . 390
11.3 A Summary of Communication . . . . . 392
11.4 The Advantages and Disadvantages of Vibrational and Acoustic Communication . . . . . 399
11.5 The Influence of the Environment on Acoustic and Vibrational Communication . . . . . 400
11.6 Information Content or the Meaning of Signals . . . . . 403
11.7 Comparing Human Language to Nonhuman Auditory Communication . . . . . 410
11.8 Summary . . . . . 412
References . . . . . 413

**12 Echolocation in Bats, Odontocetes, Birds, and Insectivores** . . 419
Signe M. M. Brinkløv, Lasse Jakobsen, and Lee A. Miller
12.1 Introduction . . . . . 419
12.2 Characteristics of Echolocation Signals . . . . . 420
12.3 Differences in Echolocation Signals in Air and Water . . . 420
12.4 Echolocation in Bats . . . . . 423
12.5 Echolocation in Odontocetes . . . . . 431
12.6 Echolocation in Birds . . . . . 440
12.7 Orientation and Echolocation in Insectivores and Rodents . . . . . 444
12.8 Are Echolocation Signals also Used for Communication? . . . . . 446
12.9 Summary . . . . . 449
12.10 Additional Resources . . . . . 450
References . . . . . 450

**13 The Effects of Noise on Animals** . . . . . 459
Christine Erbe, Micheal L. Dent, William L. Gannon, Robert D. McCauley, Heinrich Römer, Brandon L. Southall, Amanda L. Stansbury, Angela S. Stoeger, and Jeanette A. Thomas
13.1 Introduction . . . . . 459
13.2 Behavioral Options in a Noisy Environment . . . . . 465
13.3 Physiological Effects . . . . . 467
13.4 Noise Effects on Marine Invertebrates . . . . . 467
13.5 Noise Effects on Terrestrial Invertebrates . . . . . 471
13.6 Noise Effects on Reptiles . . . . . 473
13.7 Noise Effects on Amphibians . . . . . 474

13.8 Noise Effects on Fish . . . . . 477
13.9 Noise Effects on Birds . . . . . 480
13.10 Noise Effects on Terrestrial Mammals . . . . . 484
13.11 Noise Effects on Marine Mammals . . . . . 488
13.12 Summary . . . . . 492
References . . . . . 493

**Index** . . . . . 507

# About the Editor

**Christine Erbe** holds an M.Sc. degree in Physics (University of Dortmund, Germany) and a Ph.D. in Geophysics (University of British Columbia, Canada). She worked as a Research Scientist at Fisheries and Oceans Canada, was Director of JASCO Applied Sciences Australia, and after a brief stint in high-school education, returned to academia as Director of the Centre for Marine Science and Technology at Curtin University (Perth, WA, Australia). Christine's interests are underwater sound (biotic, abiotic, and anthropogenic), sound propagation, signal processing, and noise effects on marine fauna. She is a Fellow of the Acoustical Society of America, former Chair of the Animal Bioacoustics Technical Committee of the Acoustical Society of America, and former Chair of the international conference series on *The Effects of Noise on Aquatic Life*.

**Jeanette A. Thomas** (deceased), obtained her Ph.D. in Ecology and Evolutionary Biology from the University of Minnesota (1979) on underwater vocalizations of Weddell seals in the Antarctic. She was Director of the Bioacoustics Laboratory at Hubbs-SeaWorld Research Institute (San Diego, CA, USA), Senior Scientist at the Naval Ocean Systems Center (Kailua, HI, USA), and Professor in Biology at Western Illinois University (WIU; Macomb, IL, USA), where she helped establish a master's degree program in biology in collaboration with Shedd Aquarium (Chicago, IL, USA). In 2000, she developed the WIU Graduate Certificate in Zoo and Aquarium Studies. Jeanette received several awards through WIU: Distinguished Faculty Lecturer, Outstanding Researcher, and Distinguished Alumni. Jeanette was President of the Society for Marine Mammalogy (1994–1996) and Editor for Aquatic Mammals (2000–2009).

# History of Sound Recording and Analysis Equipment

1

Gianni Pavan, Gregory Budney, Holger Klinck, Hervé Glotin, Dena J. Clink, and Jeanette A. Thomas

## 1.1 Introduction

For centuries, scientists have recognized the importance of documenting human, animal, and environmental sounds. However, in recent decades, the field of bioacoustics has experienced an exceptional period of growth, primarily boosted by the rapid development of new technologies and methods to record and analyze acoustic signals. The most significant revolution in the field was the introduction of digital recording, data storage, and analysis technologies that reached the consumer market around 1980 with the introduction of the compact disc (CD). In the "analog days," researchers had to carry bulky and heavy equipment and batteries to field locations; recording duration was often limited by excessive tape and battery consumption.

Researchers produced hardcopies of sound displays using a Kay Sona-Graph™ machine and spliced together sonograms to generate figures for publication. Initially, frequency and time measurements were taken from these hardcopies using a regular ruler, and signals or sound events of interest were identified manually by listening human observers. As a result, studies using bioacoustics-based approaches were sparse. Now, researchers struggle to keep up with the ever-increasing number of studies using bioacoustics made possible by the accessibility, affordability, and extended recording capabilities of current equipment.

This chapter is a compilation of the authors' collective experiences in the field of bioacoustics, with each author having considerable experience studying the sounds of vocal animals across a myriad of terrestrial and aquatic environments. Even considering the drawbacks of the "good old days" of bioacoustics research, the authors concur they were incredibly fortunate to have a career studying fascinating animal sounds. As recording and analysis technologies improved, the types of information that could be extracted from recordings of animal sounds increased. Presently, species-level identification is possible in

Jeanette A. Thomas (deceased) contributed to this chapter while at the Department of Biological Sciences, Western Illinois University-Quad Cities, Moline, IL, USA

G. Pavan (✉)
Interdisciplinary Center for Bioacoustics and Environmental Research, Dept. of Earth and Environment Sciences, University of Pavia, Pavia, Italy
e-mail: gianni.pavan@unipv.it

G. Budney
Macaulay Library, Cornell Laboratory of Ornithology, Cornell University, Ithaca, NY, USA
e-mail: gfb3@cornell.edu

H. Klinck · D. J. Clink
K. Lisa Yang Center for Conservation Bioacoustics, Cornell Laboratory of Ornithology, Cornell University, Ithaca, NY, USA
e-mail: holger.klinck@cornell.edu; dena.clink@cornell.edu

H. Glotin
Université de Toulon, Aix Marseille Univ, CNRS, LIS, DYNI, Marseille, France
e-mail: glotin@univ-tln.fr

© The Author(s) 2022

C. Erbe, J. A. Thomas (eds.), *Exploring Animal Behavior Through Sound: Volume 1*,
https://doi.org/10.1007/978-3-030-97540-1_1

most cases, and depending on the focal animals the age, sex, reproductive status, behavior, activity patterns, and even health of an individual may be estimated from acoustic recordings. Acoustic data can be used to estimate the population density of vocal animals, and dialects can indicate the geographic boundaries of a population. However, density estimation by acoustics is still in its infancy, and will require further advancement in the spatial analysis of the acoustic environment by using multiple sensors to become reliable and widely applicable. At the community level, the entire acoustic environment or soundscape can be used to estimate species abundance and biodiversity. Changes in vocal behavior can be indicative of environmental stressors, such as anthropogenic noise or habitat degradation (Pavan 2017).

Originally, sounds of terrestrial animals were studied with equipment and methods developed for military needs, human speech analysis, and music processing (Koenig et al. 1946; Potter et al. 1947; Marler 1955). Later, scientists became interested in the sounds of aquatic animals, and underwater research was facilitated by technologies used by the navies to monitor the noise made by ships and submarines. Because of the frequency limitations of transducers (i.e., microphones and hydrophones), recorders, and analysis equipment, most initial bioacoustic research was conducted in the sonic range (i.e., the frequency range audible to humans: 20 Hz–20 kHz). Even in the early stages of the digital revolution, both recorders and analysis equipment were generally limited to audible frequencies.

A major hurdle for collecting field recordings was the large size and weight of early analog equipment, along with high power consumption, which resulted in limited recording time. The development of smaller, lightweight recording devices made the collection of acoustic data significantly easier. Currently, with the advent of small digital recorders with large solid-state memories, anyone including researchers, professionals, and amateurs can collect large amounts of high-quality acoustic data continuously over extended periods. However, when using handheld recorders, the potential influence of the human observer on the animals' acoustic behavior is a concern. Through the development and use of autonomous recorders, video cameras, and acoustic animal tags, human observer effects can be minimized, and unsupervised data collection over extended periods (days to months) and in remote locations is now possible.

In this chapter, we describe the history of the development of transducers, recorders, and sound analyzers, along with the advances that these developments facilitated in the field of bioacoustics. Recording equipment can now capture a wide range of frequencies, from infrasounds to ultrasounds (sounds below and above the range of human hearing, respectively), and are used in a wide range of applications, from the study of individuals and populations to entire soundscapes. The digital revolution in sound recording and analysis allowed for significant advances in the field of bioacoustics (Obrist et al. 2010) and resulted in the development of new disciplines, such as computational bioacoustics (Frommolt et al. 2008), acoustic ecology, soundscape ecology (Pijanowski et al. 2011a, b; Farina 2014), and ecoacoustics (Farina and Gage 2017). An overview of acoustic principles and the evolution of sound recording systems for musical applications is given in Rumsey and McCormick (2009) and in Rossing (2007).

## 1.2 Advances in Recorders

The most significant advancement in recording technology was the switch from analog-to-digital devices. A reduction in size and weight of the recorder, extended battery life, rechargeable batteries, more stable and larger capacity storage media, broader frequency range, and accessibility of a computer interface accompanied this transition. Together, these advances provided bioacousticians with an adaptable system for recording a variety of species, greater field portability, and generally more affordable high-quality equipment.

To understand the basic differences between analog and digital recorders, a clear explanation of the terms is necessary. Humans perceive the world in analog; this means that everything is

seen and heard as a continuous flow of information. In contrast, digital information estimates analog data by taking samples at discrete intervals and describing the sample values as a finite number represented by binary coding (Pohlmann 1995). For instance, while a vinyl record player (phonograph) is analog, a CD player is digital. A phonograph converts groove modulation from a vinyl record into a continuous electrical signal, whereas a CD player reads a pit structure that is interpreted as a series of ones and zeros (bits) that is typical of binary coding. Likewise, a video cassette recorder (VCR) is analog, yet a digital videodisc (DVD) player is digital. A VCR reads audio and video data from a tape as a continuous variation of magnetic information, whereas a DVD player reads ones and zeros from a disc similar to a CD.

Digital devices can approximate analog audio or video signals with an accuracy level that is dependent on both sampling rate and bit depth (or the number of bits in each sample). The Shannon-Nyquist sampling theorem proves that, for a given frequency range, a sampling rate at least twice that of the highest frequency can capture all information in that frequency band, enabling perfect reconstruction of the analog waveform.

With proper sampling, analog signals can be transformed in the digital domain at a level that makes them indistinguishable from the original. A significant advantage of digital data is that it can be stored and manipulated more easily than analog recordings. With analog recorders, each copy produces a little degradation that accumulates through multiple successive copies. Analog tapes are also prone to degradation with time. Digital copies are a perfect duplication that is indistinguishable from the original, unless specific data codes are added to identify them. More importantly, digital recordings can be directly transferred to a computer for processing or transferred through the Internet to be shared among different laboratories. If researchers want to transfer audio or video files from old analog tapes so they can be recognized and processed by a computer, they must use a sound interface based on an analog-to-digital converter (AD-converter) to digitize the analog signal and transform it into a sequence of numbers.[1] For playing back sounds from a computer, a sound interface with a digital-to-analog converter (DA-converter) is required. Next, we outline a brief history of the evolution of analog and digital recording devices. For more detail on digital recording technologies, see Pohlmann 1995.

### 1.2.1 Analog Recorders

The first purported sound recording was made by Édouard-Léon Scott de Martinville and dates back to 1860. The recording was just a few seconds in duration and was made using a phonautograph. The phonautograph has a vibrating stylus, which moves on soot-covered paper to draw the sound waveform.[2] It was invented in 1857, and although it could record sounds, it never evolved to allow reproduction of the recorded sound.

In the 1870s, Thomas Edison invented the wax-cylinder recorder (Figs. 1.1 and 1.2), which had a vibrating diaphragm that was mechanically linked to a needle that sculpted grooves. It was initially recorded on aluminum foil and then on a wax layer covering the cylinder, as it was slowly rotated and translated on a screw axis. This device encoded the sound vibrations into modulations of the groove and then allowed playback of the recorded vibrations through the same needle-membrane system.

According to Ranft (2001), the first known recordings of animal sounds (a caged Indian bird, the Common Shama) were made in Germany in 1889 on an Edison wax-cylinder. One of the first known scientific studies of animal sounds occurred in 1892 when Richard Lynch Garner recorded primates on vax cylinders at a zoo in the USA (Garner 1892). Garner also

[1] Analog Definition and Meaning: www.webopedia.com/TERM/A/analog.html; accessed 24 Oct 2021.

[2] The Phonautograms of Édouard-Léon Scott de Martinville: http://www.firstsounds.org/sounds/scott.php; accessed 24 Oct 2021.



**Fig. 1.1** Thomas Alva Edison and his phonograph. Image source: https://commons.wikimedia.org/wiki/File:Edison_and_phonograph_edit2.jpg, by Levin C. Handy (per http://loc.gov/pictures/resource/cwpbh.04044/), public domain, Wikimedia Commons

experimented with the playback of the recordings to observe the primates' reactions.

The first flat disc was invented in the late 1870s, which provided an advantage over previous technology as the discs could be easily replicated. Then in 1887, Emile Berliner patented a variant of the phonograph, named the gramophone, which used flat discs instead of spinning cylinders (Fig. 1.3). Sounds were recorded on a disc as modulated grooves, with a system similar to the one developed by Edison for wax-cylinders. The first published recording of a bird sound was issued in 1910 in Germany, and the first radio broadcast of a singing bird was in Britain in 1927 (Ranft 2001).

Lademar Poulsen, a Danish engineer, invented the telegraphone or wire recorder in 1898 (Poulsen 1900). Wire recorders were the first magnetic recording devices, and they utilized a thin metallic wire, which passed across an electromagnetic recording head. Each point along the wire was magnetized based on the intensity and polarity of the signal in the recording head. Wire recorders often had problems with kinks in the





**Fig. 1.2** Photographs of an Edison's wax-cylinder player (left) and a wax-cylinder recording (right). Image sources: (left) https://commons.wikimedia.org/wiki/File:EdisonPhonograph.jpg, by Norman Bruderhofer, www.cylinder.de, CC BY-SA 3.0 http://creativecommons.org/licenses/by-sa/3.0/, via Wikimedia Commons; (right) https://commons.wikimedia.org/wiki/File:Bettini_1890s_brown_wax_cylinder.jpg, by Jalal Gerald Aro, CC BY-SA 2.0 https://creativecommons.org/licenses/by-sa/2.0, via Wikimedia Commons

wires, but editing was relatively easy as sections of wire could simply be cut out.

In the early 1900s, RCA Victor developed the Victrola, which played records or albums that were readily available to the general public. Sounds were recorded as modulated grooves on a disc, and this disc was used to produce a master metallic plate where the grooves appeared as ridges. Albums were then produced for distribution by molding copies using the master plate and Bakelite (or synthetic plastic) material. In 1920, AT&T invented the Vitaphone, which recorded and reproduced sounds as optical soundtracks on photographic film; the film impression was made with a thin beam of light modulated by the sound.

Arthur Allen, the founder of Cornell University's Laboratory of Ornithology, and Peter Kellogg made the first recordings of wild birds in 1929 at a city park in Ithaca, NY, USA. Albert R. Brand (a graduate student of Allen) and M. Peter Keane built the first equipment for recording in the field. Together, they recorded over 40 bird species within the first two years. With World War I parabola molds available from the Physics Department, Keane and True McLean (a professor in Electrical Engineering at Cornell) constructed a parabolic reflector to improve recording of bird songs in the field[3] (Ranft 2001). In those years, Theodore Case of Fox Case Corporation approached Arthur Allen to record singing wild birds and demonstrate the sound-synchronized film technology. Under the guidance of Allen, a Fox Case Corporation crew filmed and recorded the songs of wild birds in North America (Little 2003). Today, two of those recordings can be heard on the Macaulay Library website.[4] After a successful campaign with the Fox Case film crew, Allen and his colleague Peter Paul Kellogg recorded the sounds of wildlife for research and education purposes. The Library of Natural Sounds (now known as the Macaulay Library) began in 1930 at the Cornell Laboratory of Ornithology. In 1932, Allen and Kellogg used visual and audio recordings to demonstrate to the American Ornithological Union that the ruffed grouse (*Bonasa umbellus*) produced drumming sounds (Little 2003). In 1935, Cornell biologists

[3] Macaulay Library: Early milestones (1920–1950): https://www.macaulaylibrary.org/about/history/early-milestones/; accessed 24 Oct 2021.

[4] Macaulay Library: listen to recordings of Rose-breasted Grosbeak https://macaulaylibrary.org/asset/16968 and a Song Sparrow https://macaulaylibrary.org/asset/16737; accessed 11 Oct 2021.

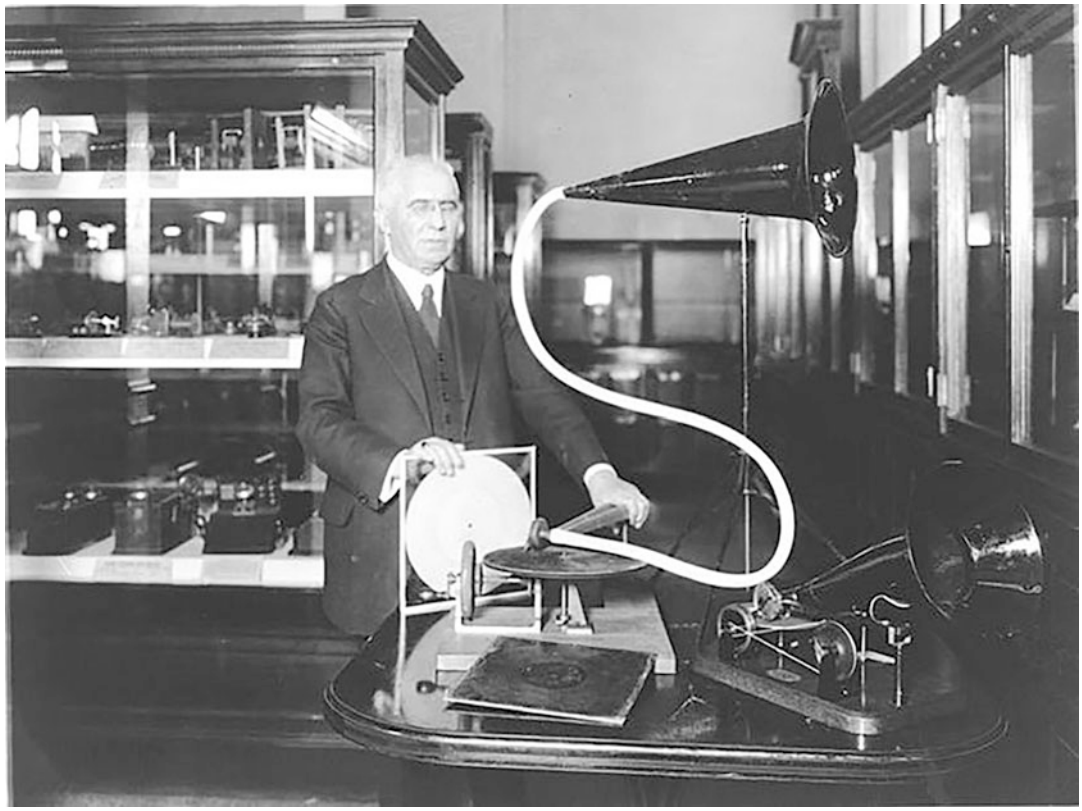


**Fig. 1.3** Emile Berliner with disc record gramophone – between 1910 and 1929. Image source: https://commons.wikimedia.org/wiki/File:Emile_Berliner_with_disc_record_gramophone_-_between_1910_and_1929.jpg, National Photo Company Collection (Library of Congress), public domain, via Wikimedia Commons

carried out an expedition to record the sounds of vanishing bird species, including the ivory-billed woodpecker (*Campephilus principalis*), for which they used a mule-drawn wagon to transport recording equipment into the field (Fig. 1.4).[5] Even with limited space and harsh conditions, Alton Lindsay, in 1934, took a phonograph recorder on the Little America Expedition to Antarctica and made recordings of airborne sounds from Weddell seals (*Leptonychotes weddellii*), available today at the Smithsonian Institution.

In the late 1930s, a German company invented the Magnetaphone, which was based on the same principle as the magnetic wire recorder, but instead of wire, it had long, thin strips of paper impregnated with fine particles of iron oxide that were drawn across an electromagnetic head. After World War II, the American company Ampex perfected the German technology by replacing paper with a thin plastic film. For almost 50 years, reel-to-reel magnetic tape was the standard media for use on recorder/playback devices (Fig. 1.5). Reel-to-reel recorders (or open-reel recorders) used variable tape speeds to record different frequency ranges, with faster recording speeds providing higher-frequency recordings. Another American company, a contemporary of Ampex, the Amplifier Corporation of America, was one of the first companies to develop a truly portable reel-to-reel recorder, the Magnemite 610, which was introduced in 1951 and was

[5] Macaulay Library: listen to the ivory-billed woodpecker recording made with an optical film recorder https://macaulaylibrary.org/asset/6784; accessed 11 Oct 2021.



**Fig. 1.4** Photograph of ornithologist Peter Paul Kellogg in 1935 in a mule-drawn wagon used to haul an amplifier (center) and optical film recorder (on the right) to capture the sounds of ivory-billed woodpeckers in the Singer Tract, Madison Parish, Louisiana. Image by Arthur A. Allen courtesy of the Cornell Laboratory of Ornithology



**Fig. 1.5** Open-reel recorder made by AEG (1939). Image source: https://commons.wikimedia.org/wiki/File:AEG_Magnetophon_K4_1939.jpg, by Friedrich Engel, CC BY-SA 3.0 https://creativecommons.org/licenses/by-sa/3.0, via Wikimedia Commons



**Fig. 1.6** Photograph of an early 1950s field recording system. Peter Paul Kellogg with an Amplifier Corporation of America Magnemite 610 reel-to-reel tape recorder and a Western Electric 633 microphone mounted in a parabolic reflector. Courtesy of the Cornell Laboratory of Ornithology

used by many pioneers in the field of bioacoustics. Figure 1.6 shows Peter Paul Kellogg using a 1950s Magnemite 610 recorder with a Western Electric 633 microphone mounted in a parabolic reflector.

Initially, tape recordings were mono recordings with one soundtrack on the tape. Stereo recording techniques (providing two record/playback channels) were developed in the 1960s. Initially, these recorders were bulky and not field portable. Then, portable open-reel recorders were developed for the rapidly developing outdoor recording needs of the radio, music, and film industries. Stereophonic recorders allowed the recording of two synchronous signals on parallel tracks onto one tape. In bioacoustics applications, often one track was used by the recordist for comments and the second track for recording animal sounds.

In the 1970s and 1980s, the most common reel-to-reel recorders used by bioacousticians were the Nagra III and IV series and the Uher 4000 series. They offered multiple recording and playback speeds (depending on the models, 3.75, 7.5, 15, or 30 inches per second), were relatively lightweight, ruggedized, and battery powered, which meant they were better suited for field studies. Eventually, recorders had even more channels (as many as 24 in some music-recording studios), which enabled scientists to record and playback signals simultaneously from more than one acoustic sensor.

Recorders were also developed to record a wide range of frequencies. Studies by Griffin (1944), Sales and Pye (1974), and Au (1993), provided evidence that animals (bats and dolphins) produce a wide range of ultrasonic signals. The first recordings of ultrasonic echolocation signals from bats and dolphins were made on expensive dedicated tape recorders at very fast tape speed (60 and 120 inches per second). Among them, the RACAL Store4DS recorder was used in the 1980s and 1990s, and it provided tape speed up to 60 inches per second to record frequencies up to 300 kHz. It was battery powered and reasonably portable. However, the limited data storage capacity of these magnetic reels meant that the recordings lasted only a few minutes.

In 1964, Philips introduced the compact cassette tape, which was comprised of a small plastic case holding two small reels with 1/8-inch wide



**Fig. 1.7** Left: Photograph of a semi-professional stereo cassette recorder Marantz CP430 used by nature recordists until the last decade of the twentieth century. Right: Photograph of a mono cassette recorder (Philips K7, 1968) with microphone and cassette inside. Image source: https://commons.wikimedia.org/wiki/File:Philips_EL3302.jpg, by mib18 at German Wikipedia, CC BY-SA 3.0 http://creativecommons.org/licenses/by-sa/3.0/, via Wikimedia Commons

magnetic tape running at 4.75 cm/s (1.875 inches per second). In the 1970s, analog cassette recorders, which could easily record and playback sounds, became available at affordable prices, but were used primarily for music and human speech, and were thus limited in frequency to the human hearing range. These recorders (Fig. 1.7) were much smaller and less expensive than reel-to-reel devices. Cassette tapes could record up to one hour on each side of the cassette (typical total recording duration was either 60, 90, or 120 min), but tapes were very thin and fragile, which made them prone to print-through (the magnetic transfer of a recorded signal to adjacent layers of tape). In 1976, Sony introduced, with little success, the Elcaset, a bigger cassette with 1/4-inch tape running at 9.5 cm/s. Today, however, it is almost impossible to find new reel-to-reel or cassette tapes as there are very few manufacturers of these media.

One of the advantages of tape recording was the possibility to play back the tapes at a speed lower or higher than the original recording speed. This way it was possible to lower the frequency of recorded ultrasonic signals to the human hearing range, thus making them audible (and longer in duration); conversely, recordings of infrasounds were played at higher speed to make them audible (and shorter in duration). The same trick can now be done easily with digital systems. Playbacks are a commonly used experimental approach in bioacoustics, wherein previously recorded sounds are broadcast to the animals of interest. Many playback studies used magnetic tape recordings containing animal sounds as the stimuli.

Researchers could easily play the sound backward (by reversing the reading direction of a spliced tape) or insert a section of tape containing sounds of another species, individual, or noise as a control stimulus. Magnetic tape was also used to record live video images. The first practical video tape recorder (VTR) was built in 1956 by Ampex Corporation. The first VTRs were reel-to-reel recorders used in television studios, which made recording for television cheaper and easier.

VHS tape recorders, introduced in the 1970s, were the first compact analog devices to record both audio and video signals simultaneously on the same tape. Commercial video cameras quickly became available for home use. Battery power for cassette recorders and VHS cameras/recorders made this equipment popular for field studies of animal behavior and sounds.

Many magnetic analog recordings had problems because the media deteriorated when tapes were not stored under properly climate-controlled

conditions. Unfortunately, some older analog recordings have been lost, or, in some cases, the players are not available to retrieve the recorded sounds. In the last decades, a great effort was made by major sound libraries to preserve old recordings (on wax-cylinders, discs, magnetic tapes, and cassettes) and to transfer them to safer digital storage (Ranft 1997, 2001, 2004). This was often not an easy task because magnetic tape recordings used a large variety of tape types, speeds, and track format arrangements. Unfortunately, many valuable tape recordings have yet to be converted to a digital format and archived. Without a long-term preservation strategy and support, it is possible that these media may be lost forever.

### 1.2.2 Digital Recorders

The introduction of the CD by the music industry in 1983 brought digital audio to the consumer market and started a new audio recording age (Pohlmann 1995). The ability to store sound in a digital format greatly improved acoustic data collection. It allowed easy and perfect replication of recordings, enabled accurate digital editing, and provided the means of more permanent data storage with direct access for processing and analysis by a computer.

In 1987, Rotary Digital Audio Tape (R-DAT or DAT) recorders were the first widely available digital recorders (Fig. 1.8). However, these devices still recorded on a thin magnetic tape encapsulated in a small cassette using a rotating helical-scanning magnetic head, which allowed for much faster head-tape speed and data density. Many R-DAT recorders allowed recording at different sampling rates of 32.0, 44.1, or 48.0 kHz and 16-bit resolution (the CD standard is 44.1 kHz, 16 bit) (Pohlmann 1995). The R-DAT format had little success in the consumer market because of the high cost but was used widely by professional recordists as a replacement for expensive and bulky open-reel recorders.

Some specialized R-DAT models allowed recording up to 100 kHz on a single channel (i.e., by using a 204.8 kHz sampling frequency and doubled tape speed). R-DAT offered recording quality that was comparable to open-reel recorders, however, the helical-scanning head proved problematic in humid conditions, and the thin tape used in R-DAT cassettes was easily damaged. An alternative to R-DAT was the digital compact cassette (DCC) introduced by Philips in 1992. DCC was compatible with the already existing analog cassette tapes but failed to gain commercial success.

Digital recorders with optical discs (CD-R and DVD-R) never gained popularity for field applications because the equipment had to remain stationary while recording. Also, at the same time, magnetic discs (hard drives) quickly became the state-of-the-art data storage media. In contrast, the MiniDisc (MD), a small optical disc developed and marketed by Sony in 1992, had more success among nature recordists,

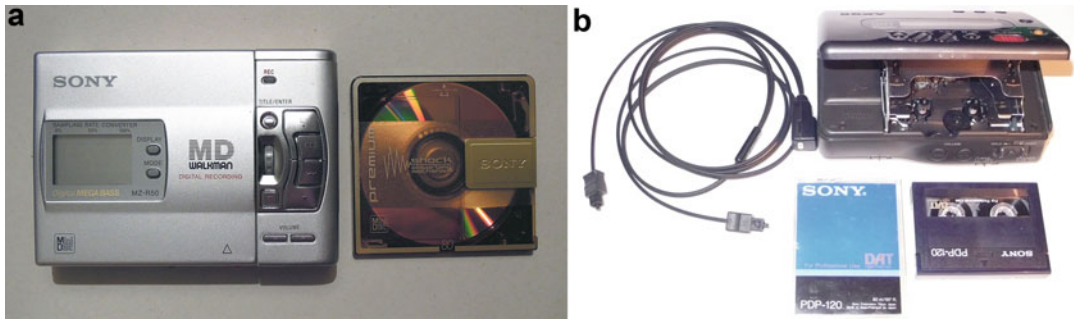



**Fig. 1.8** (**a**) Photograph of a portable R-DAT recorder Sony TCD-D7 (1992) with a DAT cassette and the optical able to provide digital data transfer to a PC. (**b**) a MiniDisc recorder and disc (1997)

because the MD portable recorders were smaller, lighter weight, and much cheaper than DAT recorders. MD offered random access to the recordings (DAT and analog tape recorders allowed only sequential access), which made it much easier to find and listen to specific sections of a recording. These devices used the same sampling mode as the CD (44.1 kHz, 16 bit). The main disadvantage of the MD was the lossy signal compression based on Adaptive Transform Acoustic Coding (ATRAC), similar to the MP3 codec developed by the Moving Picture Expert Group (Budney and Grotke 1997). The compression fit 74 minutes of acoustic data onto a small digital disc with a nominal capacity of 140 megabytes (MB) with a compression rate of 5:1. The precision of some measurements of the acoustic structure of animal sounds can be significantly affected by lossy data compression schemes (Araya-Salas et al. 2017).

With hard drive recorders and the subsequent development of solid-state memory recorders, a new generation of high-quality equipment with unparalleled capacity became available in the early 2000s (Figs. 1.9 and 1.10). Solid-state memory recorders do not require mechanical moving parts for the storage and retrieval of digital information and instead use memory cards, such as Compact Flash (CF) or Secure Digital (SD and microSD) cards also used in the digital photography market.

The subsequent development of pocket digital recorders for the consumer market allowed scientists and amateurs to record many hours of sounds with high quality. Portability and storage space increased while cost decreased. Today, tape recorders have been completely replaced by solid-state digital recorders with either external (Fig. 1.9a) or built-in microphones (Fig. 1.9c). Attempts to develop portable digital recorders based on handheld portable computers or pocket PCs never gained much popularity because of the rapid development of pocket recorders. Professional and semi-professional recorders (Fig. 1.9a) provide phantom powering at 48 V (P48) for professional condenser microphones, have quiet microphone preamplifiers, several types of powering options and can have up to 8 channels. Most pocket recorders lack the phantom powering required for professional microphones, but can power external microphones at low voltage (Plug-In-Power, or PIP; see Sect. 1.3.1).

Most digital recorders can sample at different sampling frequencies (e.g., 44.1, 48, 96, and 192 kHz) with either 16 or 24 bits of resolution, yielding very high sound quality. Some models can sample up to 192 kHz, but some of these have input electronics that limit the bandwidth to less than 60 kHz, well beyond human hearing limits, but not enough for recording animal ultrasounds. In the music industry, other standards have been developed to allow even higher acoustic quality (Melchior 2019), up to 384 kHz sampling with 32-bit depth, but they are not yet available in low-cost consumer recorders.

### 1.2.3 Recording to a Computer

In the 1990s, the first sound-acquisition boards for personal computers became available, which revolutionized the way scientists collect and analyze acoustic data. Once a sound was recorded in a digital format, recordings could easily and without degradation be transferred to a computer, stored, edited, copied, distributed, played, processed, and analyzed with different algorithms. Software (either freeware or commercial) that can be used on a laptop provides scientists with "a bioacoustics laboratory in a bag." The consumer and professional market offer a large number of sound interfaces, to be connected by USB or other standards to a PC, which can offer very high audio quality and multiple input/output channels. Smaller versions of such a setup, or compact single-board computers costing few tens of US dollars, are being used in autonomous stationary and mobile recording systems, which allow data collection and real-time data processing in remote areas for months at a time (e.g., Klinck et al. 2012).




**Fig. 1.9** (**a**) Photograph of a professional portable high-quality recorder (Sound Devices, SD722) with both hard disc and solid-state memory recording capabilities, connected to two low noise microphones (Rode NT1A) for soundscape recording. (**b**) Photograph of SONY TC-510 open-reel recorder (1982) and a SONY PCM-M10 digital recorder with its microSD memory card. (**c**) Photograph of five widely used digital recorders lined-up for comparative testing. From left: Sony PCM-M10, Sony PCM-D50, Olympus LS-3, Roland R05, and Zoom H1. They feature internal microphones, but also can connect to external Plug-In-Power (PIP) microphones or hydrophones. Courtesy of M Pesente (2016)

### 1.2.4 Autonomous Programmable Recorders

Researchers soon realized that their presence during recordings could influence the animal's behavior, and that a remote system, which could be used in the absence of human observers, was needed. There was also an increasing interest in collecting samples of the acoustic environment over long periods of time. To address these new interests, off-the-shelf recorders were modified and connected to timers, enabling recording at a defined schedule. The use of portable computers also allowed scheduled recording in the field (Fig. 1.10). However, the main limitation was the need of external batteries, which allowed only a few days of operation. In addition, long-term recording required protection of the equipment in waterproof cases and additional batteries. Defense and research laboratories alike have



**Fig. 1.10** Left: Photograph of a portable digital recording and analysis system composed of a pair of microphones, an AD-converter with USB interface (Edirol UA25), a low-power notebook, and an additional battery (2004). Right: Photograph of an autonomous terrestrial recorder by Wildlife Acoustics (model SM3, 2014) with external battery deployed in a nature reserve in Italy

interesting stories to tell about the evolution of their autonomous recording equipment (e.g., McCauley et al. 2017).

The first commercially available, programmable autonomous recorder, SongMeter 1 (SM1), was sold by Wildlife Acoustics in late 2007 and opened a rapidly developing market. Since then, new products have been proposed by companies and research groups, with increasing performances and autonomy. These can be programmed to record at defined intervals (e.g., every day across the dawn and dusk periods) or more regular sampling schedules (e.g., 1 minute every 10 minutes, or 10 minutes every half-hour) to sample temporal patterns of variation in a soundscape. This way, the acoustic behavior of animals of interest can be recorded without disturbance by the recordist and for extended periods, both day and night. These recorders need to be rugged and reliable to be deployed in harsh environments. The period of time that recorders can collect data depends on the combination of available battery power and memory. Depending on these factors, terrestrial recorders can operate for weeks to months. A grid of autonomous recorders can be used for monitoring biodiversity over a large area (e.g., entire countries; Obrist et al. 2010), even in the ultrasonic range. Figure 1.10b illustrates one type of autonomous recording system made by Wildlife Acoustics. A few different types of autonomous recorders are currently available. However, as interest in continuous, long-term acoustic monitoring of remote areas (Pavan et al. 2015; Righini and Pavan 2019) increases, new devices will continue to appear on the market and in the open-source arena. In some cases, audio recorders can be coupled with photo- and video traps to get images of the animals if they are at a close enough range.

Recent open-source autonomous recorders are built around the Raspberry Pi and similar small board computers. However, these devices often have inefficient power optimization and require large batteries to supply power over long periods. The Solo acoustic monitoring platform[6] (consisting of Raspberry Pi plus external microphone) needs a 12-V car battery to record for 40 days. Autonomous recorders need to be low-power to allow for extended periods of recording time with a manageable battery supply. The AudioMoth[7] is an open-source device that also can be purchased assembled, and it employs a low-power microcontroller with an onboard Micro Electro-Mechanical System (MEMS)

[6] Project website: https://solo-system.github.io/home.html; accessed 1 Oct 2021.

[7] https://www.openacousticdevices.info/audiomoth; accessed 22 Jun 2022.

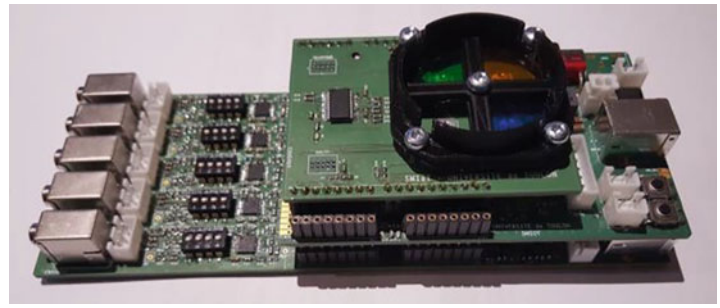


**Fig. 1.11** The JASON Qualilife also hosts a high dynamic luxmeter in four different wavelengths and direct USB HDD or micro SD storage

microphone (Hill et al. 2018). MEMS are very small and cheap and allow for production of autonomous recording devices at very low cost. Autonomous recorders can also be built around a wireless interface to send raw or processed data in real-time, in near real-time, or at scheduled intervals. However, data transmission requires power and the creation or use of a suitable wireless network (Sethi et al. 2018).

Smartphones with an external battery supply are another option used to explore animal sounds and soundscapes. The Automated Remote Biodiversity Monitoring Network (RFCx ARBIMON) can receive acoustic data from a remote recorder based on a cellphone that, if coverage is available, directly sends data to the central server with online access.[8] This system, coupled with Artificial Intelligence recognition algorithms, can identify sound categories to generate alerts to prevent poaching and deforestation. More information on autonomous recorders is available in Chap. 2.

### 1.2.5 Multi-Channel Recorders

Collecting multiple channels of acoustic data allows for acoustic localization of the sound source. Multi-channel recordings can help mitigate the Lloyd's mirror effect, a phenomenon in which low-frequency sounds near the ground may not be recorded correctly because of the interference of direct and surface reflected sound. Increased interest in collecting multiple channels of acoustic data coupled with environmental information has driven the development of new multi-channel, multi-parametric instrumentation. Multi-channel portable recorders and computer interfaces developed primarily for professional music recording can be used for bioacoustics applications, however, dedicated recorders with very high sampling rates are also being developed for specific study systems.

The recently developed JASON Qualilife[9] can record up to 5 data channels, with the maximum sampling frequency up to 800 kHz per channel, all featuring 16-bit resolution, a sharp filter to prevent aliasing, and an adjustable analog gain for a large range of uses (Fig. 1.11).

Although already designed for low-power consumption (12 V, 100 mA), to further reduce power consumption and achieve extended long-term recording, an extension board (Qualilife Wake-Up Detector; Fourniol et al. 2018; Glotin et al. 2018), can be used to trigger the recorder when it receives a signal at a specified frequency. This allows for a reduction in power consumption and data storage, also reducing unnecessary post-processing work. Moreover, it includes a high dynamic luxmeter (which works from sun zenith to lunar eclipse) that is synchronized with the acoustic recorder.

## 1.3 Advances in Microphones

There were several early attempts in the mid- to late-1800s by Johann Philipp Reis and Elisha

[8] Project website: https://rfcx.org/ & https://arbimon.rfcx.org; accessed 1 Oct 2021.

[9] Project website: https://www.univ-tln.fr/SMIoT.html; accessed 20 Jun 2022.

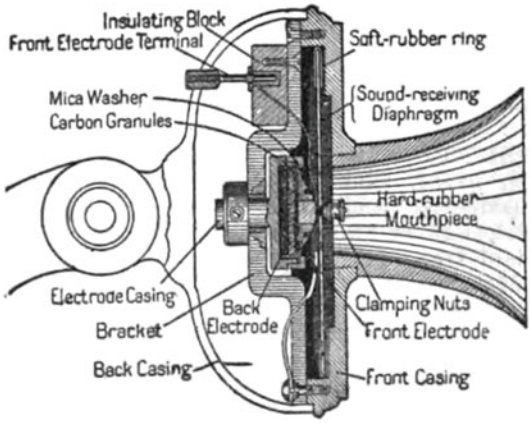





**Fig. 1.12** Left: Drawing of a carbon-button microphone (1916). Image source: https://commons.wikimedia.org/wiki/File:Carbon_button_microphone_1916.png; unknown author, public domain, via Wikimedia Commons. Right: Sennheiser MKH416 directional microphone used for bioacoustics research; https://commons.wikimedia.org/wiki/File:Sennheiser_MKH416.jpg by Galak76, CC BY-SA 3.0 http://creativecommons.org/licenses/by-sa/3.0/, via Wikimedia Commons

Gray to develop the precursor to a microphone. Reis developed the sound transmitter, which contained a metallic strip that rested on a membrane that caused intermittent contact between a metal point on the strip and an electrical circuit when it vibrated. Elisha Gray developed the liquid transmitter, consisting of a diaphragm connected to a moveable conductive rod, which was immersed in an acidic solution. In 1876, Alexander Graham Bell invented the magnetic transmitter, and Edison and Berliner developed a loosely-packed carbon granules microphone (Fig. 1.12). David Edward Hughes coined the term "microphone" in 1878 for his microphone system based on carbon granules, which performed poorly by today's standards (due to high self-noise and distortion). However, it was an important step forward, enabling technology for long-distance voice communication or telephony (for more details see Robjohns 2010)[10]

In 1886, Thomas Alva Edison refined the carbon granule microphone and developed the carbon-button transmitter. This transmitter consisted of a compartment filled with granules of carbonized anthracite coal, which were confined between two electrodes. One electrode was connected to an iron diaphragm. Edison's transmitter was durable, efficient, simple, and cheap to build. His transmitter became the basis for millions of telephone transmitters used around the world.

### 1.3.1 Microphones Used in Bioacoustics Research

At the beginning of the twentieth century, most microphones were carbon granule sensors. These early microphones were noisy and had limited sensitivity and frequency response. This meant these early microphones were suited only for recording human voices. In those early stages, dynamic microphones based on a membrane with a coil immersed in a magnetic field were difficult to produce because they required small but strong magnets.

In 1917, Edward Wente made a great stride forward by inventing the condenser microphone, which is still used in a wide variety of applications today. In the 1920s, with the significant increase in broadcast radio, there was a high demand for better quality microphones. The

[10] A Brief History of Microphones: http://microphone-data.com/media/filestore/articles/History-10.pdf; accessed 11 Oct 2021.



**Fig. 1.13** Photograph of the PRIMO EM172 microphone capsule (left) used by many nature sound recordists for their custom-made microphones (center and right). Courtesy of M Pesente

piezoelectric microphone was created based on piezoelectric crystals, which are sensitive to pressure changes and generate a voltage when compressed/decompressed; conversely, they vibrate and produce sound waves if excited by an electric signal. Originally, they used quartz or Rochelle salt crystals, but the sound quality was poor. With the development of strong magnets, dynamic microphones were then used for decades because of their simplicity and reliability. However, for bioacoustics studies, they were not sensitive enough, and their frequency response generally did not extend beyond the human hearing range. Today, almost 90% of the microphones manufactured annually are electret condenser microphones (Rossing 2007) because of their many advantages when compared with dynamic microphones, including higher sensitivity, higher fidelity, and wider frequency response. Piezoelectric transducers are now mainly used in hydrophones that have specialized ceramics that provide high sound quality. Robjohns (2010) provides a history of microphone evolution and outlines how advances in broadcast radio, telephones, television, and music industry, along with the need for directional and ultrasonic recordings, drove the design of several new types of microphones (e.g., the condenser-, dynamic-, ribbon-, and carbon-microphones).

The widely used condenser microphones are fairly sensitive, compared with dynamic microphones, and feature an extended frequency response, but they require external power. Professional condenser microphones are often powered through the signal cables with 48 V (phantom power, P48) provided by the recording device, by a preamplifier, or by a power unit. Consumer microphones usually use electret condenser capsules that require 3–5 Vdc powering (plug-in power, PIP) provided by the recorder via the microphone plug. Microphones well-suited for bioacoustics studies can be built with electret condenser capsules costing only a few US dollars (Fig. 1.13). For a detailed discussion of features and operation of microphones, see Chap. 2, section on selecting a microphone.

Many animals including insects, frogs, bats, and other terrestrial and marine mammals emit ultrasonic sounds (Sales and Pye 1974). Studies of ultrasonic signals require a broadband microphone capable of responding to signals at very high frequencies. In contrast, some animals, such as elephants, produce very low-frequency sounds and require infrasonic microphones capable of detecting signals at or below 20 Hz (Payne et al. 1986). Previously, ultrasonic and infrasonic recording required very expensive and complex transducers, recorders, and analyzers. With the

advent of broadband AD-converters in laptops and smartphones, ultrasonic and infrasonic animal sounds can now be recorded at a reasonable cost. Ultrasonic microphones may use small electret condenser capsules or MEMS, which are primarily used in smartphones. MEMS are small and inexpensive, feature an extended frequency response (including the ultrasonic frequency range), can include an AD-converter, and can be directly integrated into digital systems. Some microphones also incorporate a high-speed AD-converter and USB interface to be directly connected to a computer, a smartphone, or a tablet for recording and real-time display. The Dodotronic Ultramic series offers a range of USB ultrasonic microphones with sampling frequencies ranging from 192 kHz to 384 kHz (Buzzetti et al. 2020); the most advanced models also include the ability to record on an internal microSD memory card.[11]

In cases where researchers want to separate sounds coming from different directions, or target an individual animal for recording, a directional microphone, a parabolic reflector, or a microphone array can be used. One of the first documented attempts was in 1932, when Peter Paul Kellogg and Arthur Allen used a microphone installed in the focus of a parabolic reflector to record bird sounds (Wahlstrom 1985; Ranft 2001). Parabolic reflectors have been widely used to record animal sounds, capture distant speech, and detect the noise of incoming vehicles and airplanes during the first and second world wars (i.e., before the invention of radar; see Chap. 2 for a discussion of use and features of parabolic reflectors). As an alternative to parabolic reflectors, ultra-directional microphones, or so-called shotgun microphones, were developed. The design of shotgun microphones is based on the interference tube principle to attenuate off-axis sounds; these microphones were developed to have a narrow angle of forward reception. The shotgun was initially designed for use in a studio setting (as opposed to recording long-distance sounds) to minimize off-axis sounds (e.g., noise from the public and room reflections).

Single microphone (i.e., monophonic) recordings cannot provide any spatial information. These recordings are made with a single microphone that can be an omnidirectional microphone to capture all sounds around or a directional one to capture sounds from a specific source or direction. However, microphones can be paired to record sounds in stereo to provide a spatial sound image wherein listeners can identify the perceived spatial location of the sound source. Many different types of microphone configurations have been developed, mainly for recording music, but also for recording soundscapes.

A further development, mainly conceived for cinema and videogames, is the surround system that is based on multi-microphone (i.e., microphone array) recordings and speakers placed around the listener to create a more immersive acoustic experience (Streicher and Everest 1998; Rayburn 2011). With 3D audio, a whole acoustic space is recorded with a microphone array. From this, it is possible to extract sound information to build a stereophonic or binaural or surround program. Today 3D audio is mainly used for 3D Virtual Reality, with either video game, cinema or scientific uses, that allows the user to be placed in a 3D audio and video environment (with special visors and headphones, or in special VR rooms) and to move inside it to look and listen in any direction. The currently most used 3D audio system is Ambisonics (Fig. 1.14) that is based on 4 (first order), 8 (second order), 16 (third order) or more channels (Zotter and Frank 2019).

Specific microphone array applications in bioacoustics include localizing sound sources, either static or moving, such as flying bats (Blumstein et al. 2011). Using specific algorithms, signals can be extracted from the microphone array, and the direction and intensity of sound sources can be identified by superimposing a sound map on top of an image taken by a video camera. This type of application is called an acoustic camera and is largely employed by the automotive industry to locate sources of noise in a vehicle.

[11] Dodotronic webpage: http://www.dodotronic.com; accessed 20 Jun 2022.



**Fig. 1.14** Ambisonic recorder with 4 microphones (first order) Zoom H3VR

Acoustic cameras help visualize patterns of both indoor and outdoor noise (e.g., of a passing car, train, airplane, or around a wind turbine). Acoustic cameras have the potential to help in localizing biotic sound sources; however, they are expensive and have been rarely used for bioacoustics studies; an example is given by Stoeger et al. (2012) to identify the sound sources in elephants.

### 1.3.2 Measurement Microphones

Measurement microphones are a special class of microphones designed to make accurate amplitude measures of sounds, ranging from infrasound to ultrasound. Although measurement microphones can be used for recording, they are generally used to characterize the acoustic properties of a signal or of a location. Usually, measurement microphones are condenser microphones optimized for a specific frequency range and used to characterize a sound field or a sound level when connected to a sound level meter (or phonometer); see Chap. 2 for a discussion of measurement microphone features and operation. This microphone technology has not changed much over time; however, the measuring equipment to which microphones are connected has evolved within a few decades from bulky and expensive analog devices to small, powerful, and flexible digital devices also able to provide spectral analysis.

### 1.3.3 Accelerometers

An accelerometer measures the acceleration (i.e., the rate of change of velocity) of an object. Single- and multi-axis accelerometers can detect both the magnitude and the direction of the acceleration, as a vector quantity. They can thus measure the movements of an animal (e.g., mounted in a collar) or to sense the vibration of a body part. Tiny accelerometers are used to detect vibrations generated by insects and other animals for communication. The recently defined science of biotremology uses accelerometers and laser vibrometers to study vibrational communication in insects and other zoological groups (Hill et al. 2019) by either detecting their movements or the vibrations transmitted through the substrate. MEMS accelerometers are now very tiny and largely used in electronic devices, such as smartphones and game controllers, to sense their movement in space.

### 1.3.4 Laser and Optical Microphones

Laser microphones, also known as laser interferometers, laser accelerometers or vibrometers, are designed to detect vibrations on a surface without any contact with the sound source. These microphones can detect vibrations over large distances, from few centimeters to tens

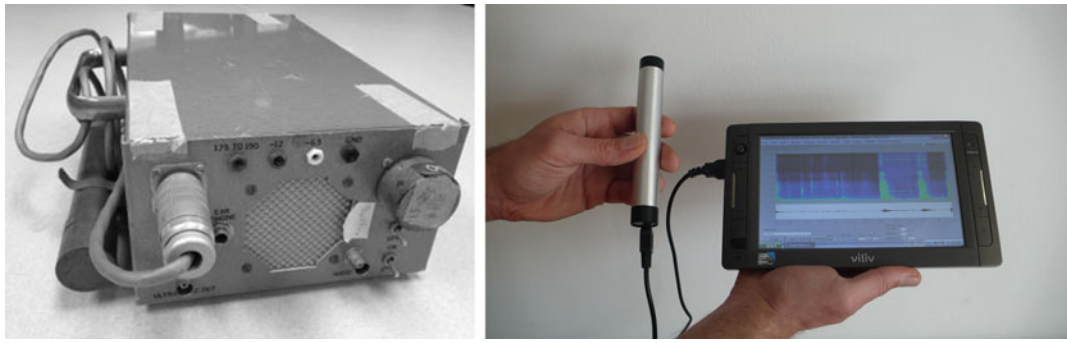


**Fig. 1.15** Left: Photograph of an early ultrasonic bat detector from the laboratory of Donald Griffin. Image courtesy of the Cornell Laboratory of Ornithology. Right: Photograph of an ultrasonic USB microphone UltraMic250k, based on MEMS, developed by Dodotronic in 2010, connected to a tablet computer that allows recording and display of ultrasounds in real-time

and hundreds of meters. For example, laser microphones can measure the vibration of a glass window to capture the sounds produced inside a room. These devices were developed for spying purposes and are now mostly used in industry to record vibration of machinery. In bioacoustics research, and biotremology studies in particular (Hill et al. 2019), this technology is used to record the vibration of animal body parts (e.g., wings or abdomen of insects producing sounds) or vibration of the substrates (e.g., plant stem, tree trunk, spider-web, and burrow-wall), which could indicate the presence of an animal. Current instruments are lightweight and easy to use; however, they require that the target being recorded is not moving and on a stable platform. These devices should not be confused with optical microphones and hydrophones, which are being developed and have a completely optical chain, where the transducer directly produces an optical signal to be sent on an optical fiber cable, either analog or digital, from the transducer to the recorder.

### 1.3.5 Bat Detectors

In the eighteenth century, the Italian scientist Lazzaro Spallanzani recognized that bats were capable of navigating and capturing their prey in the dark. While Spallanzani hypothesized that this was related to their hearing, it was not until the development of ultrasonic recorders and microphones in the early 1940s (Fig. 1.15) that scientists were able to study the ultrasonic sounds produced by bats for echolocation (Griffin 1944). Donald Griffin was working with piezoelectric transducers connected to an oscilloscope when he observed high-frequency signals produced by bats flying outside his open laboratory window. This discovery opened an entirely new field of bat echolocation research.

Early bat detectors were based on the heterodyne principle and on frequency-division counters (Obrist et al. 2010), which produced audible but highly distorted sounds when receiving ultrasonic calls. Heterodyne detectors allowed only a narrow frequency range up to a few kHz, to be shifted down to the audible range. The user then tuned the detector to the frequency of interest and listened to and recorded signals only around the tuned frequency. Information outside that frequency range was discarded.

Frequency division (or count-down) detectors cover a broad frequency range. They are based on zero-crossing detection. They count how many times the signal waveform crosses zero pressure and they produce a synthetic wave every $n$ incoming waves. The output signal frequency is a fraction of the original frequency (i.e., $1/n$), and advanced systems retain the amplitude envelope of the original signal. The frequency division

method is much better than the heterodyne; however, both produce a distorted signal often not useful for scientific investigation. The first digital models, called time-expansion detectors, digitally recorded the incoming bat calls at a high sampling rate, and played them back at a reduced sampling rate, which allowed for human observers to hear the calls and record them on a conventional recorder (Obrist et al. 2010). This method preserves all acoustic features so that recordings can be used for scientific analysis.

Digital bat detectors include a built-in ultrasonic microphone, onboard signal sampling and processing, memory for digital data storage, a graphical display to show a spectrogram with related settings, and a speaker for monitoring incoming ultrasounds by either slowing down or shifting them in frequency. Current models are completely digital, they record and store data continuously, and can transpose ultrasounds into audible sounds in real-time by spectral shifting (or spectral compression), using a Fast Fourier Transform (FFT) algorithm (see Chap. 4 on signal processing). Some bat detectors can be used as autonomous recorders which can selectively record ultrasounds from echolocating bats for many consecutive nights, with a programmable timer to start at sunset and stop at sunrise. Some also have analysis software that identifies the species, of course with variable margin of error depending on the species (see Chap. 2, section on bat detectors). Given the computing and storage capabilities of current tablets and smartphones, dedicated ultrasonic microphones with an integrated AD interface also are available to record bat calls and display their features on the device screen (Fig. 1.15).

## 1.4 Advances in Hydrophones

In 1826, Jean-Daniel Colladon and Charles-Francois Sturm made an experiment in Lake Geneva, Switzerland, to determine the speed of sound in water (Colladon 1893). They used two small boats on opposite sides of the lake, ~14 km apart. On one boat, there was an underwater bell, which was struck at the same time that gunpowder was ignited, which resulted in a paired underwater sound and above-water gunpowder flash. The operator of the second boat used an underwater listening horn to detect the sound of the bell (Fig. 1.16). The time difference between seeing the gunpowder flash and hearing the bell allowed

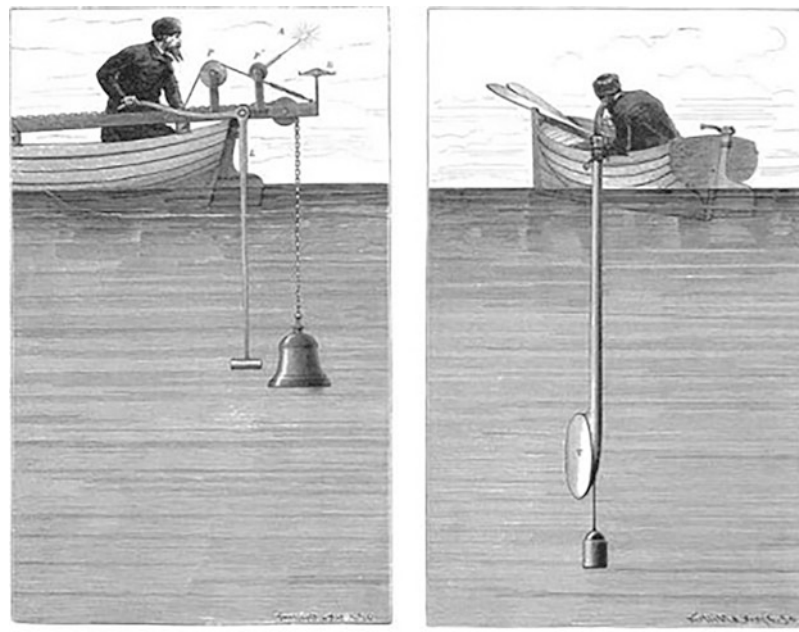


**Fig. 1.16** Experimental setup to determine the speed of sound underwater. Image Source: J. D. Colladon, Souvenirs et Memoires, Albert-Schuchardt, Geneva, 1893

the scientists to compute the speed of sound in water. Their measurements were fairly accurate and indicated that the speed of sound in water is approximately five times greater than the speed of sound in air.

Until the advent of hydrophones, it was assumed that oceans, rivers, and streams were quiet environments. Much of hydrophone development was driven by military needs during World Wars I and II, when the use of hydrophones and sonar projectors facilitated the detection of enemy vessels, particularly submarines, by listening to their sound (i.e., passive sonar) or by listening for the reflection of emitted sound pulses (i.e., active sonar). Sonar operators were some of the earliest bioacousticians who were able to distinguish sonar signals from marine animal sounds (Fish and Mowbray 1970). Today, hydrophones are used in a large variety of biological research applications to monitor population dynamics and behavior of marine invertebrates, fish, and mammals (Au and Hastings 2008; Tremblay et al. 2009). Hydrophones are also largely used to monitor the underwater noise produced by ship traffic and other invasive activities, such as seismic surveys with airguns and naval sonar (Pavan et al. 2004).

### 1.4.1 Single Hydrophones

Hydrophones are transducers used to receive underwater sound; they are usually based on piezoelectric materials. Hydrophones are generally built with a piezoelectric transducer that generates a voltage when compressed/decompressed; conversely, it can vibrate and produce sound waves if excited by an electric signal. Piezoelectric transducers can be operated either as a receiver or as a transmitter. In 1917, Paul Langevin obtained a large 10 cm × 10 cm × 1.6 cm slice of a natural quartz crystal and used this to develop a transmitter capable of emitting sound so powerful it killed nearby fish. After World War II, other materials (potassium dihydrogen phosphate, ammonium dihydrogen phosphate, and barium titanate) were used instead of quartz to build hydrophone transducers (Rossing 2007).

As the Navies of the world began to recognize the utility of listening underwater, hydrophone technology developed fairly rapidly, and also was used for oceanographic and biological research (Wenz 1962; Munk and Wunsch 1979; Urick 1983; Naramoto 2000). Most of the early bioacoustics research on aquatic animals was conducted using a battery-operated single hydrophone (Fig. 1.17) suspended in the water from the




**Fig. 1.17** Simple piezoelectric hydrophone (Aquarian Audio HC2a) with PIP powering connected to a digital pocket recorder (SONY PCM-M10)

shore, a small boat, or sea ice, and required the presence of a researcher.

Traditional hydrophones feature an analog output (voltage or current) and are available with or without a front-end preamplifier. Hydrophones that feature an integrated AD-converter and digitize the analog signal directly at the sensor are now commercially available. Some digital hydrophones also integrate signal processing and storage capabilities (e.g., real-time reporting of noise levels). Because of the increased power consumption of digital hydrophones, these are primarily used in cabled sensor networks, such as seafloor sensors or sub-surface towed arrays.

### 1.4.2 Sonobuoys

Navies of the world recognized the need for a hydrophone that could operate remotely, was mobile, and could monitor sounds at different water depths, which led to the development of sonobuoys. Sonobuoys are individual canisters that float at the water surface and house a hydrophone, dampening cable, battery, recording/transmitting electronics, and a transmitting antenna. See Chap. 2 for details of features and operation of sonobuoys. Navies of the world used sonobuoys for underwater listening to detect submarines by deploying them from airplanes or ships. A few labs were able to acquire military sonobuoys and used them for receiving and recording marine animals.

### 1.4.3 Autonomous Underwater Acoustic Recorders

In recent years, a wide variety of stationary, autonomous passive acoustic monitoring (PAM) systems have been developed for the recording of acoustic activity from naturally occurring biological and geophysical sources, as well as from anthropogenic sources in marine environments (Figs. 1.19, 1.20, 1.21, and 1.22). These systems have an advantage over systems that rely on human observers as they are non-invasive and able to collect long-term data from remote areas independently of weather and light conditions (Mellinger et al. 2007; Lammers et al. 2008; Tremblay et al. 2009; Obrist et al. 2010; Sousa-Lima et al. 2013; Jacobson et al. 2016); see Chap. 2.

### 1.4.4 Towed Hydrophone Arrays

A towed array contains several hydrophones housed in an oil-filled plastic sleeve, which are pulled behind vessels of varying size. Towed arrays of hydrophones allow beamforming (a processing technique that combines time-delayed signals from multiple hydrophones to increase gain in a given direction) to improve signal-to-noise ratio and estimate bearings to specific sound sources. Consecutive bearing estimates allow the localization of a source and determining its range. A towed array in effect provides a high-gain, directional sensor that can be steered in different directions either in real-time or in the post-processing of recordings (see Chap. 2 for details of towed hydrophone arrays). During World War I, a towed sonar array (the first documented towed array) known as the Electric Eel was developed by the US Navy physicist Harvey Hayes (Naramoto 2000). Bill Watkins and William Schevill at Woods Hole Oceanographic Institution were among the first bioacousticians to use this technology to record and study the sounds of marine mammals (e.g., Watkins and Schevill 1977; Watkins et al. 1987). The original towed arrays focused on lower-frequency signals (i.e., frequencies typical of foreign vessel noise), but Schevill and Watkins developed new instruments to record the higher frequencies emitted by dolphins. Their recordings are of high scientific value and are available online in digital format at the WHOI Watkins Sound Library.[12]

In 1983, Thomas et al. (1986, 1987) worked with a geophysical company to build a modified towed array specifically for the study of marine mammal sounds (Fig. 1.18), which was capable

[12] WHOI Library: http://cis.whoi.edu/science/B/whalesounds/index.cfm; accessed 11 Oct 2021.



**Fig. 1.18** Left: Photograph of the topside electronics required to receive, record, and process data from a towed array in 1983. Right: Photograph of deploying a towed array from the deck of a tuna seiner, the MV Queen Mary, to listen for underwater sounds of marine mammals and fish in the Eastern Tropical Pacific. Photos by Jeanette Thomas

of capturing low- and medium-frequency underwater sounds (20 Hz–15 kHz). Depth and temperature sensors on the array measured the thermocline and sound propagation conditions in the area. Self-noise from the moving ship was present, but filtered out as much as possible. Many species of marine mammals were heard, which helped the fishermen find tuna as they tend to associate with dolphin pods.

In recent years, lightweight towed arrays have been developed to meet the requirements of studying marine mammal sounds from small platforms, such as sailboats (Pavan and Borsani 1997). Deployment of the towed array from a sailboat minimizes recorded self-noise of the towing vessel. Current towed arrays can capture sounds over a large geographic area and cover a wide frequency range (from infrasound to ultrasound).

### 1.4.5 Seafloor Hydrophone Arrays

Arrays of bottom-mounted hydrophones were an important naval asset for the surveillance of oceans for the presence and movements of enemy vessels and submarines. In the 1950s, at the height of the Cold War, the US Navy launched a classified project known as the SOund SUrveillance System (SOSUS). The SOSUS large-aperture arrays allowed the Navy to detect signals at ranges of several hundred kilometers. SOSUS arrays were highly successful in detecting and tracking Soviet submarines of that era. The sailors operating the early SOSUS arrays also detected numerous biological sounds of unknown origin. An unknown low-frequency sound was attributed to the "Jezebel Monster," yet later found to be from blue (*Balaenoptera musculus*) and fin whales (*Balaenoptera physalus*). After the end of the Cold War, the SOSUS system was made available to scientists (Nishimura and Conlon 1994; Stafford et al. 1998; Watkins et al. 2000), who monitored the presence of marine mammal sounds and tracked their long-range seasonal movements across the oceans. In one case, a blue whale was tracked for 80 days along the eastern seaboard of the USA using the 20-Hz signal the animal repeatedly produced.

At present, bottom-mounted arrays of hydrophones are deployed across oceans worldwide, with some strictly dedicated to military applications, and others dedicated to monitoring



**Fig. 1.19** The JASON Qualilife DAQ 3x600 kHz in the custom array by H Glotin, recording sperm whales in the near field in 2018. Courtesy of V Sarano

earthquakes or nuclear explosions, such as the array operated by the Comprehensive Nuclear Test Ban Treaty Organization (CTBTO). Over the last decade, multidisciplinary seafloor networks were established: the North-East Pacific Time-series Undersea Networked Experiments (NEPTUNE) and the Victoria Experimental Network Under the Sea (VENUS) in Canada[13]; the Controlled, Agile, and Novel Ocean Network (CANON) run by MBARI in the USA; the European Multidisciplinary Seafloor Observatory (EMSO) run by Europe; the Submarine Multidisciplinary Observatory (SMO) managed by Italy; and the Neutrino Mediterranean Observatory (NEMO also known as KM3net) operated by the Neutrino Mediterranean Observatory. Some of these arrays are equipped with wideband hydrophones, which allow scientists to monitor a variety of marine mammal species as well as ambient noise levels (Nosengo 2009; Favali et al. 2013; Caruso et al. 2015; Sciacca et al. 2015; Viola et al. 2017). NEPTUNE and VENUS also provide online public access to recorded data. The Listening Into the Deep Ocean (LIDO) project provides real-time streaming of acoustic data that is a gateway to several underwater data acquisition systems (André et al. 2011).

[13] Canada seafloor networks: http://www.oceannetworks.ca; accessed 11 Oct 2021.

### 1.4.6 Small Arrays

Novel hydrophone array configurations have recently been developed for a team led by François Sarano to conduct a longitudinal study on the same group of sperm whales since 2013, under the authority of the Marine Megafauna Conservation Organization and as part of the global program Maubydick. In 2017 and 2018, the team collected a set of audio-visual recordings using a custom acoustic antenna developed by the University of Toulon with the JASON Qualilife DAQ (Data AcQuisition) to record the animals in the near field at very high frequency (600 kHz sampling frequency, Fig. 1.19). A similar antenna has been deployed in Amazonia allowing high-definition 3D tracking and click analysis of the Amazon river dolphin (*Inia geoffrensis;* Glotin et al. 2018).

## 1.5 Autonomous Mobile Systems

### 1.5.1 Aerial Mobile Systems

Autonomous mobile monitoring systems were developed for terrestrial applications, such as the Autonomous Aerial Acoustic Recording Systems (AAARS) developed at the University of Tennessee (Buehler et al. 2014). This system is based on an altitude-controlled weather balloon with an acoustic recorder and a GPS unit with

radio transmitter. It moves quietly according to local winds and can be tracked by a radio receiver. If ground anchored, this system allows the recording of sounds in a given location. Mobile systems based on drones, on the contrary, can be stationary or can be programmed to survey a given area, however, they are very noisy and this can severely affect animal behavior and both the quality and usability of the recordings.

### 1.5.2 Underwater Mobile Systems

The high cost of visual and acoustic marine surveys conducted from large research vessels drove the development of new monitoring solutions using autonomous vehicles; either moving on the surface (Unmanned Surface Vessels, USVs) or underwater (Autonomous Underwater Vehicles, AUVs). These systems are remotely operated by an onshore pilot and can monitor offshore areas for weeks or months at a time (Klinck et al. 2012, 2015).

The most commonly used autonomous mobile systems to monitor the marine acoustic environment are underwater gliders (Baumgartner et al. 2013). These instruments (Fig. 1.20) use small changes in buoyancy, in conjunction with wings, to convert vertical motion to horizontal motion, and thereby propel themselves forward with very low-power consumption. Gliders slowly dive (~ 0.25 m/s horizontal speed) in a saw-tooth pattern through the water. When surfacing after a dive, the glider communicates with an onshore base station to exchange data and commands (e.g., send position, remaining battery capacity, whale detections, and ambient noise levels, and receive new waypoints). The maximum operating depth of current models is about 1000 m. Therefore, these instruments are well-suited for monitoring of deep-diving odontocetes, such as beaked whales (Klinck et al. 2012).

Other instruments in this category include deep-diving (Matsumoto et al. 2013) and surface drifters (Griffiths and Barlow 2015). These instruments drift with the ocean current and cannot be programmed to navigate along a defined track-line. However, they are much cheaper than gliders. Recent Autonomous Surface Vehicles (ASV) can perform surveys along a pre-defined track; among these, the Sphyrna (Fig. 1.20) has advanced algorithms to allow 3D passive acoustic tracking of deep divers with four hydrophones fixed on the keel (Poupard et al. 2019).




**Fig. 1.20** Left: Photograph of the passive acoustic seaglider™ developed by the Applied Physics Laboratory, University of Washington. Courtesy of G Shilling. Right: The Sphyrna ASV allows 3D passive acoustic tracking of diving cetaceans

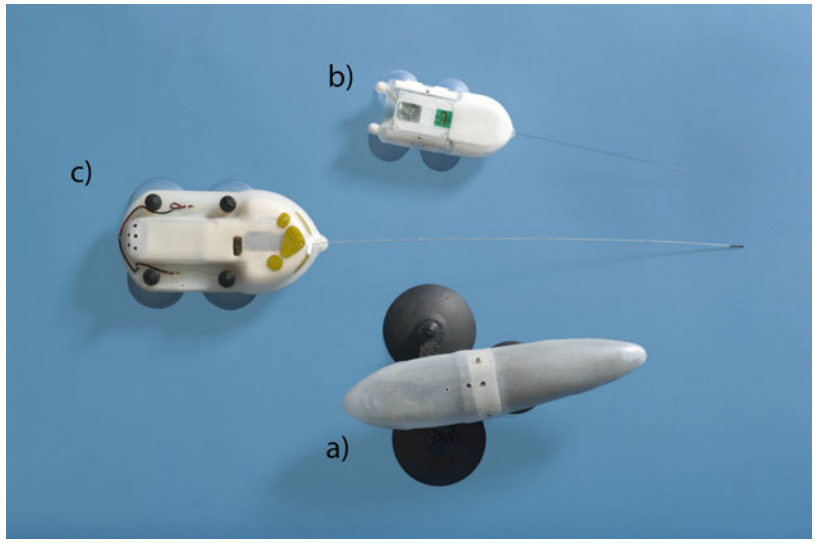



**Fig. 1.21** The evolution of the DTAG over fifteen years. Each design comprises electronics, batteries, suction cups, floatation material, and a VHF transmitter for retrieval when the tag is floating on the sea surface. The tags all record sound, depth, and motion to solid-state memory. However, the size, capabilities, and endurance have changed over the years. The earliest version developed in 2000 (a) had 400 MB of memory and could record a single sound channel at 16 kHz sampling frequency for a few hours. The most recent version developed in 2009 (b) records stereo sound at up to 500 kHz sampling frequency for almost two days. (c) is an intermediate version of the tag. Courtesy of P Tyack and M Johnson (2016)

### 1.5.3 Animal Acoustic Tags

A recent development for studying animals in-situ is the animal-worn acoustic tag. Such devices allow detailed observations of the movement and acoustic behavior of tagged animals. However, for some species, such as cetaceans, developing a reliable, long-term instrument attachment has been problematic.

Recorders in collars, similar to those used for radio tracking, have also been experimented to record sounds and activity of terrestrial animals while moving freely, but with few applications. More successful was using the crittercam developed and used by National Geographic to primarily provide amazing video[14] of wild animals either on land or in water. Lynch et al. (2013) attached an inexpensive collar-mounted recording device on ten wild mule deer (*Odocoileus hemionus*) over two weeks in Colorado. Recorded sounds included rumination, which allowed the researchers to document foraging activities.

Video tags have been attached to whales, dolphins, sirenians, and penguins, and to document the underwater life. Sophisticated acoustic tags provided an important step forward in marine mammal bioacoustics. The development of these tags was primarily driven by the need to document and understand the reaction of cetaceans to underwater sounds such as naval sonars, airguns, and pile drivers. The D-TAG (Johnson and Tyack 2003), A-Tag (Akamatsu et al. 2007), Acousonde recorder (Burgess et al. 2011), and other similar instruments, feature a variety of animal movement detectors (three-axial accelerometer, magnetometer, depth-sensor, light sensor, etc.) and acoustic sensors (hydrophones). These tags are attached to the animals with non-invasive suction cups, and usually stay attached for a few hours, but can stay on the animal for up to a few days. Once detached, the tag floats to the surface and transmits a radio signal to aid recovery. This kind of technology (Fig. 1.21) has enabled important

[14] https://www.nationalgeographic.org/education/crittercam-education/; accessed 11 Oct 2021.

research on sound usage and behavioral responses of animals to anthropogenic sounds, such as naval sonars (Tyack 2009; Tyack et al. 2011).

Often a variety of sensors can be attached to the animal to provide additional environmental or behavioral data to accompany acoustic recordings. Evans et al. (2004) attached a waterproof video camera with a hydrophone, VHS recorder, and depth-sensor to examine vocal behavior during dives of Weddell seals in Antarctica. Each time the seal vocalized, the depth and time of the sound were documented, audio and video were recorded, and the call type was later analyzed in the laboratory. Researchers had to retrieve the VHS tapes, but this species remains close to a colony during the breeding season, hauls out on the ice daily, and is easily (re)captured for recovery of the tag and data. Current digital video equipment is highly miniaturized and allows new exciting options for exploring the life of animals in the wild.

## 1.6 Advances in Sound Analysis Hard- and Software

The most important advancements in sound analysis equipment were the transition from analog-to-digital systems, along with the transition from hardware to software signal processing. This provided lightweight, field portable, battery-operated units with higher storage capacity, more stable storage media, and broadband analysis, often at a more affordable price than before. Now, even a smartphone can produce a spectrogram in real-time. Another important breakthrough was the ability of scientists to share digital data using the internet and shared storage in the cloud.

Initially, the basic analysis of acoustic signals was done using oscilloscopes. These instruments provided a visual representation of the waveform of acoustic signals known as oscillograms, which are plots with amplitude on the *y*-axis and time on the *x*-axis. Originally, oscilloscopes were large, heavy, expensive, AC powered, and used vacuum tubes. To obtain a hardcopy of the waveform, a camera was used to capture an image from the display. In some cases, the waveforms were traced on paper by an oscillating pen (similar to a seismometer).

The Kay Electric Company (later to become Kay Elemetrics) developed the Sona-Graph™ machine, which was a completely analog instrument and one of the first instruments to create an image of a sound known as a SonaGramTM. Developed primarily for navy applications and initially called vibralyzer, this technology was applied successfully to the study of human speech and animal sounds (Koenig et al. 1946; Borror and Reese 1953; Thorpe 1954; Marler 1955: Fig. 1.22). A SonaGram (sometimes called a sonogram by biologists) is a visual representation of the frequencies (on the *y*-axis) and intensity (color or shades of gray as the *z*-axis) in a sound as they vary with time (on the *x*-axis). This type of image visualization is also called spectrogram. The Sona-Graph™ was very expensive and capable of analyzing a signal of only a few seconds in duration up to 8 or 16 kHz. The device offered two analysis settings, wideband (300 Hz) and narrowband (45 Hz). The wideband setting provided better time resolution, while the narrowband setting provided better frequency resolution (Beecher 1988). The sound could be played back from a reel-to-reel recorder and recorded on an iron oxide magnetic track, which ran the circumference of a large internal turntable. A special thermo- sensitive paper was wrapped around a drum mounted on top of the turntable. The drum spun synchronously with the turntable as the signal was played back through a variable band-pass filter or a filter bank, and a stylus burned the signal onto the paper on the rotating drum according to the level of sound at the frequencies given by the filter (Fig. 1.23).

This was a smelly, smoky process, which made the procedure unpleasant for researchers. To analyze a long sound recording, several short spectrogram sections had to be printed and taped together. The resulting sheets of paper often required a lot of wall or table space for review and further analysis. Because of the large size, these spectrograms were also difficult to reduce in size and adapt for inclusion in a publication.

In the 1970s, a camera using Kodak photographic paper (the size of 35-mm film) was attached to the screen of an advanced



**Fig. 1.22** Photograph of L. Irby Davis using an early Kay Electric Co. Sona-Graph Sound Spectrograph analyzer (the late 1950s). Notice the sonogram on the paper wrapped around the drum on top of the analyzer. Courtesy of the Cornell Laboratory of Ornithology

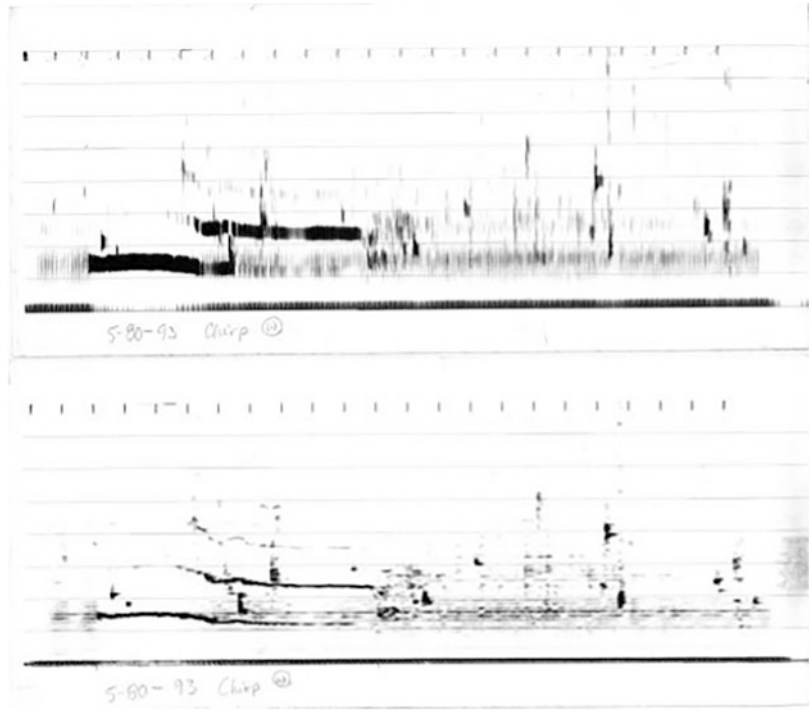



**Fig. 1.23** Two spectrograms by Ken Norris illustrating the wide-band (top) and narrow-band settings (bottom) of the Kay Sona-Graph 6061A spectrum analyzer. Note that the values of the *x*- and *y*-axes were not printed on the output. The *x*-axis is time in seconds and *y*-axis is the frequency in hertz. Courtesy of the Cornell Laboratory of Ornithology

oscilloscope capable of performing real-time FFT spectrum analysis (Hopkins et al. 1974). As the sound played, a spectrogram image appeared on the screen and the camera photographed the resulting image in real-time. Measurements of frequency and time could be taken as the spectrograms were displayed. The photographic paper had to be developed in a dark room and produced a roll of 35-mm paper about 4 m long. One advantage of this system was the ability to

view the sounds in real-time, which allowed scientists to study patterns of sounds. This system produced long-lasting spectrograms that are still usable 40 years later (see Thomas and Kuechle 1982 for samples of sonogram output).

Once thermal imaging paper (similar to the paper used in older fax machines) was developed, Kay, Unigon, and other companies developed real-time spectrogram imaging units, which had a continuous output using large rolls (8 inch wide) of thermal imaging paper. For further analysis, segments had to be cut with scissors. However, these data were difficult to analyze, store, and prepare for publication. Measurements of frequency and time could be taken as the images were displayed on the analyzer but were not provided on the output itself. If exposed to light or heat, the hardcopies gradually turned brown and were generally unusable after a few years.

In the mid-1970s, the first attempts were made to use general-purpose computers to analyze sounds, mainly for speech analysis. These attempts used the Fast Fourier Transform (Strong and Palmer 1975), an algorithm that decomposes a signal segment into a finite number of sinusoids, each one characterized by frequency, amplitude, and phase. This algorithm was successfully applied to the human voice and to animal sounds to produce spectrograms in different formats. The speed and data-handling capabilities of computers in subsequent years allowed for the implementation of more complex mathematical signal processing algorithms (see Chap. 4 on signal processing).

A few years later, in 1980, a computer-based digital spectrographic workstation was developed at the University of Pavia (Italy) that produced black-and-white spectrograms of animal sounds on a computer screen, with a moving cursor to take measures. The workstation produced and printed a spectrogram of a 1-s signal in about 40 minutes (Pavan 1983, 1985). The AD-converter allowed users to acquire and analyze sounds in the ranges of 5, 10, and 20 kHz with a sampling frequency of 51.2 kHz. Hardcopies of displays were made on the computer's printer and then joined together (Fig. 1.24).

Around that same time, in 1984, a group of acousticians at The Rockefeller University and

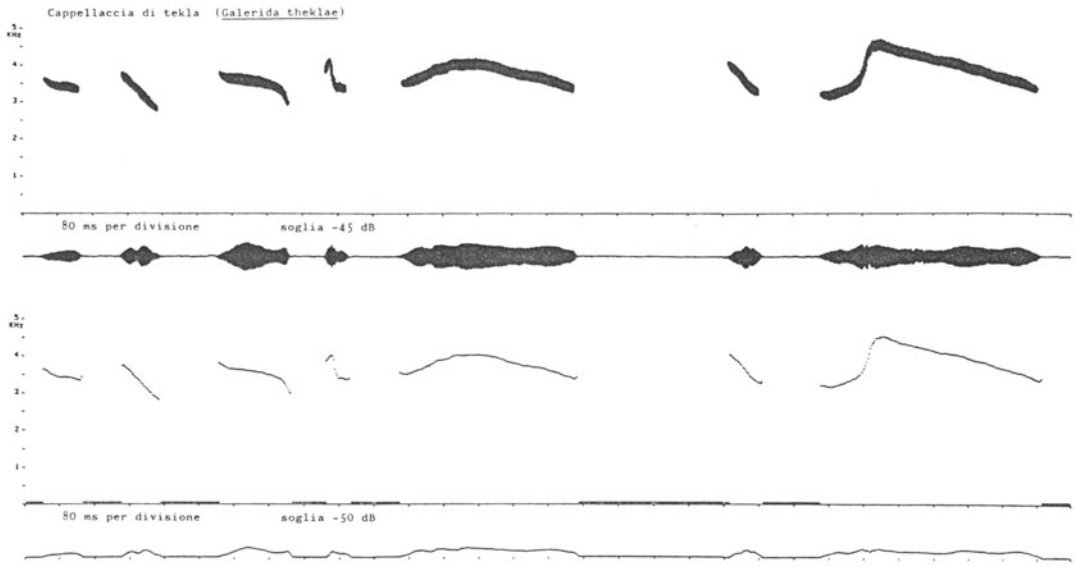



**Fig. 1.24** Black-and-white spectrogram of a 2.4-s bird song (Thekla lark) produced in 1981 by joining three printouts of 800 ms each; the spectrogram generation required 2 hours. The *x*-axis is time in seconds and *y*-axis is the frequency in hertz. Frequency range 0–5 kHz, sampling frequency 20,480 Hz, and 12-bit resolution (72-dB dynamic range). From top: spectrogram, envelope, tracking of dominant frequency, and amplitude plot in dB

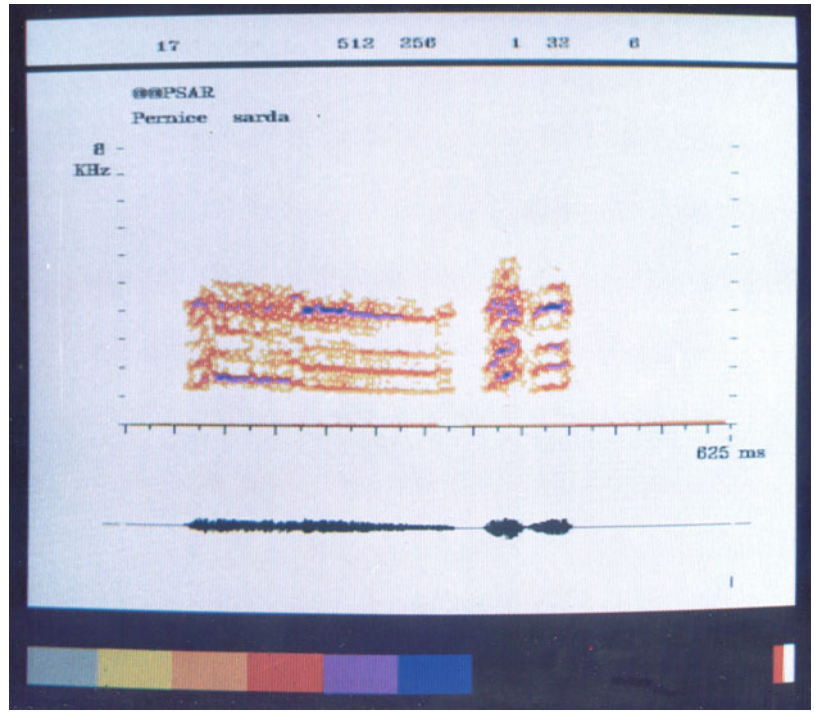



**Fig. 1.25** Photograph of an envelope-plot and color spectrogram generated by the digital signal processing workstation based on HP1000 mainframe in 1985. Recordings were of calls of a Barbary partridge (*Alectoris barbara*)

Engineering Design Inc. developed a software program, called *Signal*. This software was developed for computers and was able to control and communicate with the recording hardware. The system was able to display spectrograms in real-time, provide basic time-frequency information of recorded signals, and store data digitally on the computer's hard disc. These developments revolutionized bioacoustics sound analysis; however, at the time, these units were expensive, custom-made, and had very little storage capacity (the typical storage available in 1985 was 5 MB on a 15-inch magnetic disc).

In 1985, the spectrographic workstation was upgraded to produce color spectrograms (Fig. 1.25; Pavan 1992) on a mainframe computer (HP 1000) interfaced to an AD-converter and to a graphic workstation.[15] Around this time, the first personal computers (PC) appeared, and the software was rewritten to produce real-time color spectrograms and signal envelopes using an Intel 8086/8087 processors and a high-quality Audiologic Duetto sound board produced in Italy, with sampling frequency up to 48 kHz with 16-bit resolution, and later with a widely available and cheap Sound Blaster sound card. A mouse-driven cursor allowed to take accurate measures directly on the computer screen, and printouts were possible in gray scales on standard matrix-dot printers or on thermal printers. By storing the recordings in a digital format, it was also possible to edit the recordings and to play them back at a different speed or even backward (e.g., to produce playback tapes for behavioral experiments).

At the same time, other researchers started experimenting with digital signal processing. Aubin (France) and Specht (Germany) developed similar digital sound analysis systems that also included the synthesis of sounds for playback experiments (Bremond and Aubin 1989; Specht 1992; Aubin et al. 2000). Specialized AD-converters appeared on the market to sample analog signals at high rates, which allowed digital recording and analysis of

[15] http://www.unipv.it/cibra/res_dspwstory_uk.html; accessed 29 Oct 2021.

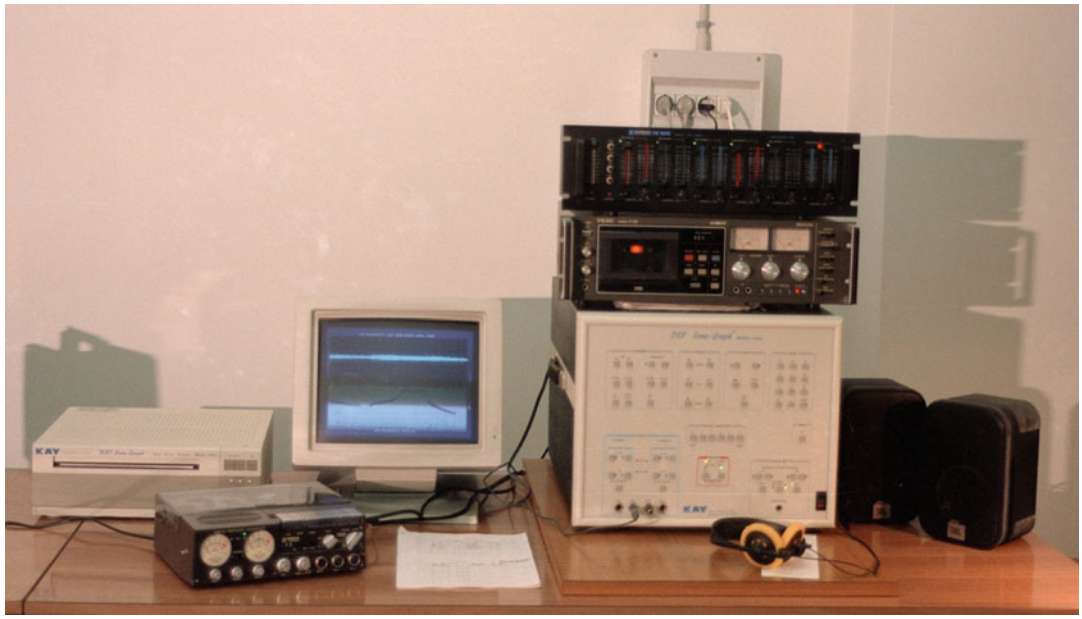


**Fig. 1.26** Photograph of the University of Pavia bioacoustic laboratory equipment in 1989 with a Kay Sona-Graph DSP 5500, color monitor, thermal printer, portable open-reel stereo recorder, cassette deck recorder, filter bank, speakers, and headphone

frequencies up to 100 kHz. However, specialized processors (Digital Signal Processors, DSP) were required to process ultrasonic signals in real-time (Pavan 1992, 1994).

In 1987, new commercially available digital instruments dedicated to sound analysis became available, among them the Kay Sona-Graph DSP 5500 (Fig. 1.26). This very expensive unit was able to analyze and display stereo signals in real-time up to 32 kHz. Either reel-to-reel or cassette recordings could be used as an input, and the unit had a thermal-paper printer for printing gray-shaded spectrograms.

Digital sound storage and analysis became widespread given the improvements in digital computer technology and data storage, coupled with the proliferation of personal computers, and the development of dedicated sound analysis software packages. These advances also fostered the development of high-quality electro-acoustic and musical equipment (microphones, recorders, and AD-converters) for a rapidly expanding consumer market of musicians and music enthusiasts. Among the first analysis software dedicated to bioacoustics, it is worth to mention *Canary*, developed for Macintosh computers at Cornell University, then replaced by *Raven,*[16] a multi-platform software developed from the same university. For an overview of computer-based bioacoustics sound analysis and related algorithms, see Hopp et al. (1998), Zimmer (2011), and Sueur (2018). Many academic institutions and companies started to develop software programs for PC, Mac, and Linux computers.[17]

These software programs allowed for easy recording, manipulation, analysis, and display of signals. Now, researchers are able to collect huge acoustic datasets, and computational bioacoustics faces the Big Data problem. The latest software programs, either commercial or open source, also enable the user to run sophisticated detection/

[16] Accessed from the K. Lisa Yang Center for Conservation Bioacoustics https://ravensoundsoftware.com/software/raven-pro/; accessed 11 Oct 2021.

[17] List of available software: http://tcabasa.org/?page_id=2666; accessed 4 Oct 2021. https://github.com/rhine3/bioacoustics-software; accessed 20 Jun 2022.

classification algorithms over long-term data sets for automated detection of occurrences of a target sound (see Chap. 8 on detection and classification methods). This saves much time and avoids having to view and listen to the entire recording manually. Scientists also can use readily available programming environments (including *MATLAB*, *Octave*, *Python*, *R*) to develop their own analyses, often facilitated by libraries of procedures dedicated to sound processing and bioacoustic analysis (e.g., Sueur et al. 2008; Sueur 2018; Ulloa et al. 2021).

In the late 1990s, smartphone technology was developed, along with sound analysis software for these devices. Smartphones of the twenty-first century have the same computing power as a desktop PC. Sound recording and visualization applications were developed for both Android and iPhone Operating System (iOS) platforms. In addition, the development of the Internet of Things and low-cost computer platforms (e.g., Arduino, Raspberry PI, and others) have allowed scientists to build web-enabled data recording and analysis systems. These new technologies and analytical methods can be applied not only to audible sound but also to infrasonic and ultrasonic signals. For example, ultrasonic echolocation signals produced by bats can now easily be shifted into the human hearing range, visualized, and analyzed in real-time with handheld digital devices, with a smartphone equipped with an ultrasonic microphone, or remotely monitored with web-connected recorders.[18]

## 1.7 Summary

Advances in electronic technology over the last 100 years, including the dramatic size reduction of equipment, increased battery life, increased data storage capacity, the switch from analog-to-digital recorders, along with the transition from analog-to-digital signal processing, have facilitated an explosion of research in the field of bioacoustics. Many of these advances were enabled by equipment developed for military use, professional music applications, human speech analysis, and for the radio, television, and film industries. Often an improvement in one type of equipment led to advancements in another. Analog devices, which stored data on magnetic tape, were replaced by digital devices, such as optical discs, hard drives and solid-state memory cards. Microphones and hydrophones are now used in arrays that allow long-term monitoring, localization of the sound-producing animals, and 3D acoustic recording. Towed hydrophone arrays allow mobile surveys of marine sounds, which can be coupled with animal sightings and environmental data. Autonomous transducer/recorder units can be deployed for long-term monitoring of biotic and abiotic sounds in both air and water in remote habitats. Recently, smartphone applications have provided an affordable and portable bioacoustics laboratory for use by hobbyists, citizen scientists, and researchers alike.

The digital revolution in sound recording and analysis has facilitated significant advances in the field of bioacoustics and enabled the development of ecoacoustics, which joins bioacoustics and ecology, and computational bioacoustics. Acousticians are now able to study the sounds from sound-producing species in a wide variety of locations, during day and night, year-round, and often remotely. Many free and commercially available software packages for recording and analyzing acoustic data have been developed for computers, tablets, and smartphones. Artificial Intelligence is now being applied to big data problems and to bioacoustic recordings to hopefully classify and recognize sounds at species level. It has never been easier or cheaper to study the acoustic world ranging from infrasounds to ultrasounds. However, it is always important to know the intrinsic limitations of each piece of equipment or software, the constraints given by the environmental context, and all their potential impact on the final results. It is also worth considering that bioacoustics and ecoacoustics are now being widely used to study and monitor critical and endangered species and to monitor entire ecosystems to understand climate change impacts.

[18] http://www.bat-pi.eu/; accessed 11 Oct 2021.

## References


Akamatsu T, Telemann J, Miller LA, Tougaard J, Dietz R, Wang D, Wang K, Siebert U, Naito Y (2007) Comparison of echolocation behaviour between coastal and reverie porpoises. Deep-Sea Res II 54(3–4):290–297

André M, van der Schaar M, Zaugg S, Houégnigan L, Sánchez AM, Castell JV (2011) Listening to the deep: live monitoring of ocean noise and cetacean acoustic signals. Mar Pollut Bull 63:18–26

Araya-Salas M, Smith-Vidaurre G, Webster M (2017) Assessing the effect of sound file compression and background noise on measures of acoustic signal structure. Bioacoustics 28(1):57–73. https://doi.org/10.1080/09524622.2017.1396498

Au WWL (1993) The Sonar of Dolphins, Springer, Berlin, 278 pp

Au WWL, Hastings MC (2008) Principles of marine bioacoustics. Springer, Berlin

Aubin T, Rybak F, Moulin B (2000) A simple method for recording low-amplitude sounds. Application to the study of the courtship song of the fruit fly *Drosophila melanogaster*. Bioacoustics 11:51–67

Baumgartner MF, Fratantoni DM, Hurst TP, Brown MW, Cole TVN, Van Parijs SM, Johnson M (2013) Real-time reporting of baleen whale passive acoustic detections from ocean gliders. J Acoust Soc Am 134: 1814–1823

Beecher MD (1988) Spectrographic analysis of animal vocalizations: implications of the "uncertainty principle". Bioacoustics 1:187–208

Blumstein DT, Mennill DJ, Clemins P, Girod L, Yao K, Patricelli G, Deppe JL, Krakauer AH, Clark CW, Cortopassi KA, Hanser SF, McCowan B, Ali AM, Kirschel ANG (2011) Acoustic monitoring in terrestrial environments using microphone arrays: applications, technological considerations and prospectus. J Appl Ecol 48:758–767

Borror DJ, Reese CR (1953) The analysis of bird songs by means of a vibralyzer. Wilson Bulletin 65(4):271–303

Bremond JC, Aubin T (1989) Choice and description of a method of sound synthesis adapted to the study of bird calls. Biol Behav 14:229–237

Budney GF, Grotke RW (1997) Techniques for audio recording vocalizations of tropical birds. In: Remsen JV (ed) Ornithological Monographs 48: studies in neotropical ornithology honoring Ted Parker. American Ornithologists' Union, Chicago, IL, pp 147–163

Buehler, DA, Hockman EV, Prevost EC, Wilkerson JB, Smith DR, Fischer RA (2014) Demonstration and implementation of autonomous aerial acoustic recording systems (AAARS) to monitor bird populations in inaccessible areas. Proceedings of AOUCOSSCO 2014, Estes Park, Colorado, USA, p 17

Burgess WC, Oleson EM, Baird RW (2011) A hydrodynamic acoustic recording tag for small cetaceans and first results from a pantropical spotted dolphin. Poster presented at the 19th Biennial Conference on the Biology of Marine Mammals, Tampa, Florida, November–December 2011

Buzzetti F, Brizio C, Pavan G (2020) Beyond the audible: wide band (0-125 kHz) field investigation on Italian Orthoptera (Insecta) songs. Biodivers J 11(2):443–496

Caruso F, Sciacca V, Bellia G, De Domenico E, Larosa G, Papale E, Pellegrino C, Pulvirenti S, Riccobene G, Simeone F, Speziale F, Viola S, Pavan G (2015) Size distribution of sperm whales acoustically identified during long term deep-sea monitoring in the Ionian Sea. PLoS One 10:e0144503. https://doi.org/10.1371/journal.pone.0144503

Colladon JD (1893) Souvenirs et memories. Albert-Schuchardt, Geneva

Evans WE, Thomas JA, Davis R (2004) Sounds used during underwater navigation by a Weddell seal (*Leptonychotes weddellii*). In: Thomas JA, Moss C, Vater V (eds) Echolocation in bats and dolphins. Univ Chicago Press, pp 541–547

Farina A (2014) Soundscape ecology. Springer:315 pp

Farina A, Gage SH (eds) (2017) Ecoacoustics: the ecological role of sounds. Wiley, Hoboken. 352 pp

Favali P, Chierici F, Marinaro G, Giovanetti G, Azzarone A, Beranzoli L, De Santis A, Embriaco D, Monna S, Lo Bue N, Sgroi T, Cianchini G, Badiali L, Qamili E, De Caro M. G, Falcone G, Montuori C., Frugoni F., Riccobene G, Sedita M., Barbagallo G, Cacopardo G, Calì C, Cocimano R, Coniglione R, Costa M, D'Amico A, Del Tevere F., Distefano C, Ferrera F, Valentina Giordano V, Massimo Imbesi M, Dario Lattuada D, Migneco E, Musumeci M, Orlando A, Papaleo R, Piattelli P, Raia G, Rovelli A, Sapienza P, Speziale F, Trovato A, Viola S, Ameli F, Bonori M, Capone A, Masullo R, Simeone F, Pignagnoli L, Zitellini N, Bruni F, Gasparoni F, Pavan G (2013) NEMO-SN1 abyssal cabled observatory in the Western Ionian Sea. IEEE J Ocean Engineer 38(2): 358–374

Fish MP, Mowbray WH (1970) Sounds of Western North Atlantic fishes. A reference file of biological underwater sounds. The John Hopkins Press, Baltimore

Fourniol M, Gies V, Barchasz V, Kussener E, Barthelemy H, Vauche R, Glotin H (2018) Low-power frequency trigger for environmental internet of things, in IEEE/ASME Mechatronic & Embedded Systems. http://sabiod.org/bib

Frommolt KH, Bardeli R, Clausen M, (editors), 2008) Computational bioacoustics for assessing biodiversity. Proceedings of the International Expert meeting on IT-based detection of bioacoustical patterns. Published by Federal Agency for Nature Conservation, Bonn, Germany

Garner RL (1892) The speech of monkeys. William Heinemeann, London

Glotin H, Blakefield G, Trone M, Bonnett DE, Giraudet P, Giès V, Barchasz V, Patris J, Malige F, Balestriero R (2018) 1 MHz SR high definition 3D tracking of wild Amazon River dolphins with JASON tiny array, results and perspectives, in detection, classification,

localisation, & density estimation workshop. Sorbonne, Paris
Griffin DR (1944) Echolocation by blind men, bats and radar. Science 29:589–590
Griffiths ET, Barlow J (2015) Design of the drifting acoustic spar buoy recorder (DASBR). Proceedings of the 7th international workshop on detection, classification, localization, and density estimation of marine mammals using passive acoustics, San Diego, CA, USA, pp. 86
Hill AP, Prince P, Covarrubias EP, Doncaster CP, Snaddon JL, Rogers A (2018) AudioMoth: evaluation of a smart open acoustic device for monitoring biodiversity and the environment. Methods Ecol Evol:1–13
Hill P, Lakes-Harlan R, Mazzoni V, Narins PM, Virant-Doberlet M, Wessel A (eds) (2019) Biotremology: studying vibrational behavior. Springer Verlag, p 534. https://doi.org/10.1007/978-3-030-22293-2
Hopkins CD, Rosetto M, Lutjen A (1974) A continuous sound Spectrum analyzer for animal sounds. Z Tierpsychol 34(3):313–320. https://doi.org/10.1111/j.1439-0310.1974.tb01804.x
Hopp SL, Owren MJ, Evans CS (eds) (1998) Animal acoustic communication: sound analysis and research methods. Springer, p 421
Jacobson EK, Forney KA, Harvey JT (2016) Evaluation of a passive acoustic monitoring network for harbor porpoise in California. Tech rep no. CEC-500-2016-008. https://doi.org/10.13140/RG.2.1.2282.9680
Johnson MP, Tyack PL (2003) A digital acoustic recording tag for measuring the response of wild marine mammals to sound. IEEE J Ocean Eng 28:3–12
Klinck H, Mellinger DK, Klinck K, Bogue NL, Luby JC, Jump WA, Shilling GS, Litchendorf T, Wood AS, Schorr GS, Baird RW (2012) Near-real-time acoustic monitoring of beaked whales and other cetaceans using a seaglider™. PLoS One 7(5):e36128. https://doi.org/10.1371/journal.pone.0036128
Klinck H, Fregosi S, Matsumoto H, Turrpin A, Mellinger DK, Erofeev A, Barth JA, Shearman RK, Jafarmadar K, Stelzer R (2015) Mobile autonomous platforms for passive-acoustic monitoring of high-frequency cetaceans. Proceedings of the 8th International Robotic Sailing Conference, Mariehamn, Åland, Finland, August 2015: 29–38
Koenig W, Dunn HK, Lacy LY (1946) The sound spectrograph. J Acoust Soc Am 18:19–49
Lammers MO, Brainard RE, Whitlow WWL, Mooney TA, Wong KB (2008) An ecological acoustic recorder (EAR) for long term monitoring of biological and anthropogenic sounds on coral reefs and other marine habitats. J Acoust Soc Am 123:1720–1728
Little RS (2003) For the birds: the Laboratory of Ornithology and Sapsucker Woods at Cornell University. Basking Ridge, NJ
Lynch E, Angeloni L, Fristrup K, Joyce D, Wittemyer G (2013) The use of on-animal acoustical recording devices for studying animal behavior. Ecol Evol 3(7): 2030–2037. https://doi.org/10.1002/ece3.608
Marler P (1955) Characteristics of some animal calls. Nature 176:6–8
Matsumoto H, Jones C, Klinck H, Mellinger DK, Dziak RP, Meinig C (2013) Tracking beaked whales with a passive acoustic profiler float. J Acoust Soc Am 133: 731–740
McCauley RD, Thomas F, Parsons MJG, Erbe C, Cato D, Duncan AJ, Gavrilov AN, Parnum IM, Salgado-Kent C (2017) Developing an underwater sound recorder. Acoust Aust 45(2):301–311. https://doi.org/10.1007/s40857-017-0113-8
Melchior VR (2019) High resolution audio: a history and perspective. J Audio Eng Soc 67(5):246–257
Mellinger DK, Stafford KM, Moore SE, Dziak RP, Matsumoto H (2007) An overview of fixed passive acoustic observation methods for cetaceans. Oceanography 20(4):36–45
Munk W, Wunsch C (1979) Ocean acoustic tomography: a scheme for large scale monitoring. Deep-Sea Res 26A:123–161
Naramoto MV (2000) A concise history of acoustics in warfare. Appl Acoust 59:137–147
Nishimura CE, Conlon DM (1994) IUSS dual use: monitoring whales and earthquakes using SOSUS, marine tech. Soc J 27:13–21
Nosengo N (2009) The neutrino and the whale. Nature 462:560–561
Obrist MK, Pavan G, Sueur J, Riede K, Llusia D, Marquez R (2010) Bioacoustic approaches in biodiversity inventories. In: manual on field recording techniques and protocols for all taxa biodiversity inventories. Abc Taxa 8:68–99
Pavan G (1983) Ricerche di elettronica applicata alla biologia. 1. Analisi computerizzata del canto degli uccelli. Pubbl Ist Entom Univ di Pavia 24:1–43
Pavan G (1985) Analisi con calcolatore delle emissioni acustiche degli uccelli. Annuario EST (Enciclopedia della Scienza e della Tecnica). Arnoldo Mondadori, Milano, pp 135–140
Pavan G (1992) A portable DSP workstation for real-time analysis of cetacean sounds in the field. European Research on Cetaceans 6:165–169
Pavan G (1994) A digital signal processing workstation for bioacoustical research. Atti 6 conv. Ital. Ornit. Torino 1991:227–234
Pavan G (2017) Fundamentals of soundscape conservation. In: Farina A, Gage SH (eds) Ecoacoustics: the ecological role of sound. Wiley, pp 235–258
Pavan G (web) http://www.unipv.it/cibra/res_dspwstory_uk.html. Accessed 28 March 2018
Pavan G, Borsani JF (1997) Bioacoustic research on cetaceans in the Mediterranean Sea. Mar Freshw Behav Physiol 30:99–123
Pavan G, Fossati C, Manghi M, Priano M (2004) Passive acoustics tools for the implementation of acoustic risk mitigation policies. In: Evans PGH, Miller LA (eds) Proceedings of the workshop on active sonar and

cetaceans, 17th ECS conference, March 2003. European Cetacean Society Newsletter no. 42 – Special Issue: 52-58
Pavan G, Favaretto A, Bovelacci B, Scaravelli D, Macchio S, Glotin H (2015) Bioacoustics and ecoacoustics applied to environmental monitoring and management. Rivista Italiana di Acustica 39(2):68–74
Payne KB, Langbauer WR Jr, Thomas EM (1986) Infrasonic calls of the Asian elephant (*Elephas maximus*). Behav Ecol Sociobiol 18:297–301
Pijanowski BC, Farina A, Gage SH et al (2011a) What is soundscape ecology? An introduction and overview of an emerging new science. Landsc Ecol 26:1213–1232. https://doi.org/10.1007/s10980-011-9600-8
Pijanowski BC, Villanueva-Rivera LJ, Dumyahn SL et al (2011b) Soundscape ecology: the science of sound in the landscape. Bioscience 61:203–216. https://doi.org/10.1525/bio.2011.61.3.6
Pohlmann KC (1995) Principles of digital audio. McGraw-Hill, New York
Potter RK, Kopp GA, Green HC (1947) The visible speech, Bell Laboratory, Cambridge, MA
Poulsen V (1900) Method of recording and reproducing sounds or signals. US Patent 661619A
Poupard M, Ferrari M, Schluter J, Marxer R, Giraudet P, Barchasz V, Gies V, Pavan G, Glotin H (2019) Real-time passive acoustic 3D tracking of deep diving cetacean by small non-uniform Mobile surface antenna. IEEE ICASSP:8251–8255
Ranft R (1997) The wildlife section of the British library National Sound Archive. Bioacoustics 7:315–319
Ranft R (2001) Capturing and preserving the sounds of nature. In: Linehan A (ed) Aural history: essays on recorded sound. The British Library, London, pp 65–78
Ranft R (2004) Natural sound archives: past, present and future. An Acad Bras Cienc 76(2):455–465
Rayburn RA (2011) Eargle's the microphone book: from mono to stereo to surround - a guide to microphone design and application, 3rd edn. Elsevier, Oxford
Righini R, Pavan G (2019) First assessment of the soundscape of the integral nature reserve "Sasso Fratino" in the central Apennine, Italy. Biodivers J 21(1):4–14. https://doi.org/10.1080/14888386.2019.1696229
Robjohns H (2010) A brief history of microphones. http://microphone-data.com/media/filestore/articles/History-10.pdf
Rossing TD (ed) (2007) Springer handbook of acoustics. Springer, New York. 1167 pp with CD-ROM
Rumsey F, McCormick T (2009) Sound and recording. Focal Press, New York. 628 pp
Sales G, Pye D (1974) Ultrasonic communication by animals. Chapman and Hall, London
Sciacca V, Caruso F, Beranzoli L, Chierici F, De Domenico E, Embriaco D, Favali P Giovanetti, Larosa G, Marinaro G, Papale E, Pavan G, Pellegrino C, Pulvirenti S, Simeone F, Viola S, Riccobene G (2015) Annual acoustic presence of fin whale (*Balaenoptera physalus*) offshore eastern Sicily, Central Mediterranean Sea, PLoS One, 10 (11): e0141838. https://doi.org/10.1371/journal.pone.0141838
Sethi SS, Ewers RM, Jones NS, Orme CDL, Picinali L (2018) Robust, real-time and autonomous monitoring of ecosystems with an open, low-cost, networked device. Methods Ecol Evol 9(12):2383–2387
Sousa-Lima R, Norris TF, Oswald JN, Fernandes DP (2013) A review and inventory of fixed autonomous recorders for passive acoustic monitoring of marine mammals. Aquat Mamm 39:23–53
Specht R (1992) Sonagraphishce Analysen mit dem personal computer. Der Falke 1992:281–283
Stafford KM, Fox CG, Clark DS (1998) Long-range acoustic detection and localization of blue whale calls in the Northeast Pacific Ocean. J Acoust Soc Am 104(6):36161–33625
Stoeger AS, Heilmann G, Zeppelzauer M, Ganswindt A, Hensman S, Charlton BD (2012) Visualizing sound emission of elephant vocalizations: evidence for two rumble production types. PLoS One 7(11):e48907. https://doi.org/10.1371/journal.pone.0048907
Streicher R, Everest FA (1998) The new stereo Soundbook. AES Publishing, Pasadena, CA
Strong WJ, Palmer EP (1975) Computer-based sound spectrograph system. J Acoust Soc Am 58:899–904
Sueur J (2018) Sound analysis and synthesis with R. Springer, New York
Sueur J, Aubin T, Simonis C (2008) Seewave: a free modular tool for sound analysis and synthesis. Bioacoustics 18:213–226
Thomas JA, Kuechle VB (1982) Quantitative analysis of Weddell seal (*Leptonychotes weddellii*) underwater vocalizations at McMurdo Sound, Antarctica. J Acoust Soc Am 72(6):1730–1738
Thomas JA, Awbrey FT, SR Fisher (1986) Use of acoustic techniques for studying whale behavior. Rep Int Whal Commn Special Issue 8, No. SC/A82/Bw10: 121–138
Thomas JA, Fisher SR, Ferm LM, Holt RS (1987) Acoustic detection of cetaceans using a towed-array of hydrophones. Rep Int Whal Commn Special Issue 8, No. SC/37/03: 139–148
Thorpe WH (1954) The process of song-learning in the chaffinch as studied by means of the sound spectrograph. Nature 173:465
Tremblay C, Calupca T, Clark CW, Robbins M, Spaulding E, Warde A, Kemp J, Newhall K (2009) Autonomous seafloor recorders and auto detection buoys to monitor whale activity for long-term and near-real-time applications. J Acoust Soc Am 125: 2548
Tyack PL (2009) Human-generated sound and marine mammals. Phys Today, November: 39–44
Tyack PL, Zimmer WMX, Moretti D, Southall BL, Claridge DE, Claridge DE, Durban JW, Clark CW, D'Amico A, DiMarzio N, Jarvis S, McCarthy E, Morrissey R, Ward J, Boyd IL (2011) Beaked whales respond to simulated and actual navy sonar. PLoS One

6(3):e17009. https://doi.org/10.1371/journal.pone.0017009
Ulloa JS, Haupert S, Latorre JF, Aubin T, Sueur J (2021) Scikit-maad: an open-source and modular toolbox for quantitative soundscape analysis in Python. Methods Ecol Evol 12(12):2334–2340. https://doi.org/10.1111/2041-210X.13711
Urick RJ (1983) Principles of underwater sound. McGraw-Hill, New York
Viola S, Grammauta R, Sciacca V, Bellia G, Beranzoli L, Buscaino G, Caruso F, Chierici F, Cuttone G, Embriaco D, Giovanetti G, Favali P, Pavan G, Pellegrino C, Pulvirenti S, Riccobene G, Simeone F (2017) Continuous monitoring of noise levels in the Gulf of Catania (Western Ionian Sea). Study of correlation with ship traffic. Mar Pollut Bull 121(1–2):97–103. https://doi.org/10.1016/j.marpolbul.2017.05.040
Wahlstrom S (1985) The parabolic reflector as an acoustical amplifier. J Audio Eng Soc 33(6):418–430
Watkins WA, Schevill WE (1977) Sperm whale codas. J Acoust Soc Am 62:1485–1490
Watkins WA, Tyack PL, Moore KE (1987) The 20 Hz signals of finback whales (*Balaenoptera physalus*). J Acoust Soc Am 82:1901–1912
Watkins WA, George GE, Daher MA, Mullin K, Martin DL, Haga SH, DiMarzio N (2000) Whale call data for the North Pacific November 1995 through July 1999. Occurrence of calling whales and source locations from SOSUS and other acoustic systems. WHOI Technical Report 2000–02: 160 pp
Wenz G (1962) Acoustic ambient noise in the ocean: spectra and sources. J Acoust Soc Am 34:1936–1956
Zimmer WMX (2011) Passive acoustic monitoring of cetaceans. Cambridge University Press, Cambridge
Zotter F, Frank M (2019) Ambisonics. Springer Verlag, Cham

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



# Choosing Equipment for Animal Bioacoustic Research

2

Shyam Madhusudhana, Gianni Pavan, Lee A. Miller, William L. Gannon, Anthony Hawkins, Christine Erbe, Jennifer A. Hamel, and Jeanette A. Thomas

## 2.1 Introduction

Until a few decades ago, progress in bioacoustic and then ecoacoustic research was severely limited by available equipment. Over time, technological advances and the availability of user-friendly analysis software have made bioacoustics research more commonplace. The advantage of passive bioacoustic studies (in which sounds are often remotely recorded) is that the methods are non-invasive and anyone with a minimal amount of equipment can record animal sounds. However, this disadvantage diminishes if a researcher is not knowledgeable about the characteristics and limitations of the equipment being used. Given the rapid advances in digital technology, bioacousticians are often challenged with keeping up with these advances. Appropriate selection and usage of sensors, amplifiers, filters, and recorders, and proper usage of analysis software are key to valid studies on animal sounds. This chapter guides bioacoustics researchers in selecting appropriate gear for maximizing the outcomes of their research.

To record, store, and play back sounds, there are two types of devices: analog and digital. Analog recording devices, such as cassette recorders and reel-to-reel tape recorders, are now obsolete and almost completely replaced by digital recording devices. However, many researchers over time have made phonograph, reel-to-reel, or cassette recordings, which provide historical data. So, when reading an older research article in bioacoustics, one may have to consider the potential limitations of the specific equipment used at the time and their ramifications on the reported findings. Chapter 1 provides an overview of older and historic equipment.

Jeanette A. Thomas (deceased) contributed to this chapter while at the Department of Biological Sciences, Western Illinois University-Quad Cities, Moline, IL, USA

S. Madhusudhana (✉)
K. Lisa Yang Center for Conservation Bioacoustics, Cornell Lab of Ornithology, Cornell University, Ithaca, NY, USA
e-mail: shyamm@cornell.edu

G. Pavan
Department of Earth and Environment Sciences, University of Pavia, Pavia, Italy
e-mail: gianni.pavan@unipv.it

L. A. Miller
Institute of Biology, University of Southern Denmark, Odense M, Denmark
e-mail: lee@biology.sdu.dk

W. L. Gannon
Department of Biology and Graduate Studies, Museum of Southwestern Biology, University of New Mexico, Albuquerque, NM, USA
e-mail: wgannon@unm.edu

A. Hawkins
The Aquatic Noise Trust, Kincraig, Blairs, Aberdeen, UK

C. Erbe
Centre for Marine Science and Technology, Curtin University, Bentley, WA, Australia
e-mail: c.erbe@curtin.edu.au

J. A. Hamel
Department of Biology, Elon University, Elon, NC, USA
e-mail: jhamel2@elon.edu

© The Author(s) 2022

C. Erbe, J. A. Thomas (eds.), *Exploring Animal Behavior Through Sound: Volume 1*,
https://doi.org/10.1007/978-3-030-97540-1_2

## 2.2 Basic Concepts of Sound Recording

The acquisition, storage, and playback of sounds in digital systems involve the interoperation of a few independent components (Fig. 2.1). Bioacoustics researchers may choose to source the necessary components and assemble a setup themselves. The practical considerations for selecting these components will be covered in Sect. 2.3. Alternatively, researchers may opt for pre-assembled equipment. The growing market has made available a wide variety of programmable, and often customizable, autonomous recorders. Section 2.4 discusses a few of the widely used terrestrial and underwater autonomous recorders. Organizations developing autonomous recorders often invest in the necessary trial-and-error experimentation for arriving at optimal combinations of components for different applications. The use of such pre-assembled equipment allows bioacoustics researchers to circumvent the associated efforts (financial and labor). However, unique demands of specific studies may not always be addressed by existing autonomous recorders. Before diving into details of each component, we provide a quick recap of the overarching concepts and terminologies.

### 2.2.1 Sampling Rate and Bandwidth

The sampling rate used when converting analog electronic signals to digital signals limits the maximum frequency that can be recorded. The sampling frequency is measured in hertz, and the sampling rate (which has the same value but different unit) is measured in samples/s. The frequency range is limited by the Nyquist frequency, which is ½ of the sampling frequency (see Chap. 4). Sampling frequency for the standard CD is 44.1 kHz (i.e., high enough to match the full human hearing range). An 8-kHz sampling frequency suffices to understand the human voice. Nowadays, digital recorders easily sample up to 192 kHz and higher, with the flexibility to choose lower sampling frequencies (32, 44.1, 48, 88.2, and 96 kHz are common). Instrumentation recorders can have sampling frequencies up to 1 MHz.

Despite the available sampling frequencies, the actual recording bandwidth of a recorder is dictated by the analog electronics before the analog-to-digital (AD) converter. Because most commercial recorders are designed for the recording of music or human speech, the upper frequency is often limited to 20 kHz and the electronics do not have a flat frequency response beyond this limit, even if selecting a high sampling frequency such as 192 kHz. For professional recorders, the real frequency response (i.e., the output amplitude across frequencies as a function of input amplitude) is usually stated in the equipment specifications (e.g., flat to within ±3 dB between 10 Hz and 60 kHz). If the frequency response is not specified, it is important to make some tests using a frequency-generator as a sound source. It is also important to consider that the frequencies close to the Nyquist frequency might be affected by artifacts such as aliasing.

### 2.2.2 Aliasing

According to sampling theory, to preserve all information in an analog signal, a sampling frequency at least twice the highest frequency in the signal (including harmonics) should be used. A

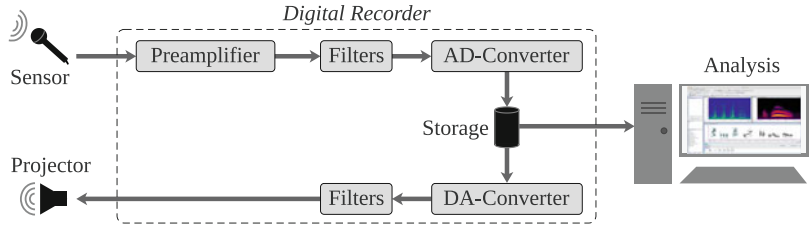



**Fig. 2.1** Signal chain of a typical digital recording setup in bioacoustics studies showing the different components involved in the collection, analysis, and transmission of sounds

non-optimal sampling frequency can produce misrepresentations of components in the original waveform, which often manifest as artifacts in a spectrographic display but are not actually present in the original signal (see Chap. 4, section on aliasing). In a spectrogram, the alias is mostly in the higher frequency region and appears as the mirror-image of the actual signals beyond the Nyquist frequency (Fig. 2.2). In digital recording, anti-aliasing filters (Sect. 2.3.2.2) are required before the sampling stage to prevent aliasing from sounds that have components higher than the Nyquist frequency.

### 2.2.3 Amplitude Sensitivity

Amplitude sensitivity, expressed as the ratio of output voltage to input pressure, indicates how many volts are produced from a sound with a root-mean-square (rms) sound pressure of 1 Pa in air and 1 μPa in water. More commonly, sensor sensitivity is given in decibel: dB re 1 V/Pa for microphones and dB re 1 V/μPa for hydrophones. To convert the linear sensitivity to dB, one needs to take 20 $\log_{10}$. So, a microphone sensitivity of 1 mV/Pa (=0.001 V/Pa) can be expressed as −60 dB re 1 V/Pa. Note that an rms sound pressure of 1 Pa is equal to a sound pressure level (SPL) of 94 dB re 20 μPa, because

$$1\ \text{Pa} = 1{,}000{,}000\ \mu\text{Pa} = 50{,}000 \times 20\ \mu\text{Pa};$$

$$\text{apply } 20 \log_{10} \text{ and get: } 20 \log_{10}(50{,}000) = 94.$$

The most sensitive sensor is not necessarily the "best" sensor. When attempting to capture very loud sound, less sensitive equipment should be chosen to avoid signal distortion or, in extreme cases, damaging the equipment. If only a sensor of low sensitivity is available, then an amplifier may be used in the recording chain, but self-noise may become an issue. High sensitivity allows lower gain settings to promote a good recording.

### 2.2.4 Bit-Resolution and Dynamic Range

The dynamic range is the difference between the highest and lowest sound levels that can be recorded. Digital recorders usually operate with 16- or 24-bit resolution; 16 bits guarantee a

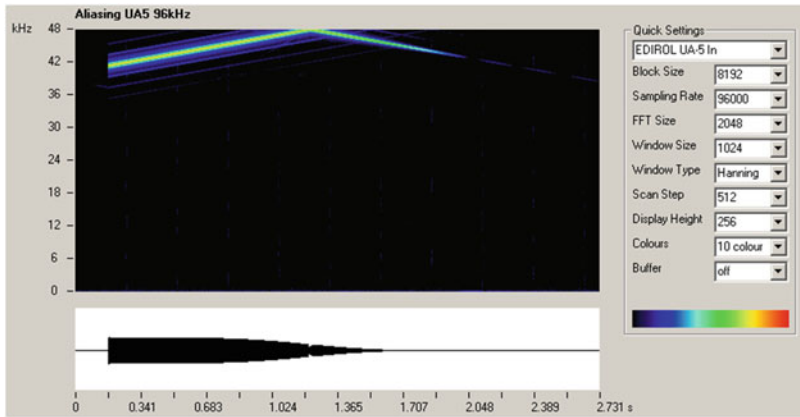



**Fig. 2.2** Spectrogram (top) and oscillogram (bottom) of an AD-converter with a sinusoidal frequency sweep from 40 kHz to 100 kHz as input. Sampling frequency 96 kHz, and thus Nyquist frequency 48 kHz. In an ideal system with a sharp anti-aliasing filter, the spectrogram would only go up to 48 kHz and show nothing once the signal frequency went beyond Nyquist. In this real-world example, however, as the signal frequency $f$ exceeds the Nyquist frequency $f_N$, the alias (appearing as the downsweep) is created with frequency $f-f_N$. As such, a 50-kHz input produces a 46-kHz alias and a 52-kHz input produces a 44-kHz alias, etc. The amplitude of the alias depends on the attenuation of the anti-aliasing filter at the input frequency. An attenuation of −10 dB at 50 kHz produces an alias at 46 kHz with a level of −10 dB relative to the input level. Spectrogram generated by *SeaPro* (http://www.unipv.it/cibra/seapro.html; accessed 15 Mar. 2021) software

dynamic range of about 96 dB (unipolar, 90 dB bipolar) and 24 bits theoretically produce a dynamic range of 144 dB (unipolar, 138 dB bipolar) thus encompassing the dynamic range of human hearing. However, even the best analog circuits rarely exceed 110 dB of dynamic range. This means that of the available 24 bits, only 20 bits are effectively used to encode the sound and the others are dominated by noise. In many conditions, the real dynamic range is limited to 70–80 dB by the noise of the sensor and preamplifier. An accurate setting of the recording levels can allow effective use of 16-bit recorders, without wasting the extra storage space required for 24-bit recording. However, when incoming sound levels cannot be predicted, the 24-bit setting allows additional dynamic range for unpredictable sound events (e.g., high-intensity impulsive noises such as from pile driving). The recorded volume should be set at a particular level to exploit the dynamic range of the recording setup: high enough to rise above the equipment self-noise during quiet times, but not too high to cause clipping of loud sounds. Recently introduced recorders allow 32-bit floating-point recording by combining the output of two 24-bit converters working with different signal gains. This simplifies the setting of recording levels but cannot yet overcome the dynamic range limitations of the microphones and of associated preamplifiers.

### 2.2.5 Self-Noise

All components of the signal chain suffer from self-noise, which is additive across the signal chain. Self-noise and dynamic range are the two critical specifications that affect amplitude response. For example, when recording in very quiet locations or to pick up very low-level sounds, the self-noise generated by the components of a signal chain must be taken into consideration, along with dynamic range. Self-noise limits the spatial range of bioacoustic sampling. It may also be an issue in playback, when self-noise is amplified and broadcast in addition to the intended signal. The circuits inside sensors can generate broadband background noise with various spectral shapes (i.e., not necessarily flat across the frequency band, like white noise, but worse at higher frequencies). The level of this noise is expressed in decibel (e.g., dB(A) after frequency weighting, dB re 20 μPa unweighted in air, or dB re 1 μPa unweighted in water) to indicate the equivalent sound level of noise as if generated by the environment. The self-noise of a sensor is almost always declared in its technical specifications; the same is true for professional recorders. On the contrary, for many consumer recorders, even of high quality, the self-noise measures are rarely available. A useful comparison of the self-noise of consumer recorders available on the market is presented on the website of Avisoft Bioacoustics.[1]

The noisiest component of the chain determines the quality of the recording. This is particularly important when recording low-level sounds (Fig. 2.3). The input self-noise is expressed as the Equivalent Input Noise (EIN) measured in an open or unloaded circuit and expressed in dBU (the "U" stands for "unloaded"). Very good values range from −130 dBU to −120 dBU, and poor recorders have a −100 dBU EIN.

## 2.3 Instrumentation of Signal Chain Components

To ensure that proper equipment is used for recording, analysis, and playback, researchers must consult manuals for each piece of equipment in the signal chain before conducting research. In some cases, laboratory tests may be required to verify the real performance or to calibrate equipment (Sect. 2.6). While recording, researchers must ensure that the frequency response (and, in turn, bandwidth), self-noise, and dynamic range (in particular, the maximum recording level) of the overall recording system do not end up deleting or significantly distorting a portion of the signal. Otherwise, a researcher can miss part of

[1] http://www.avisoft.com/recorder-tests/; accessed 1 Feb. 2021.

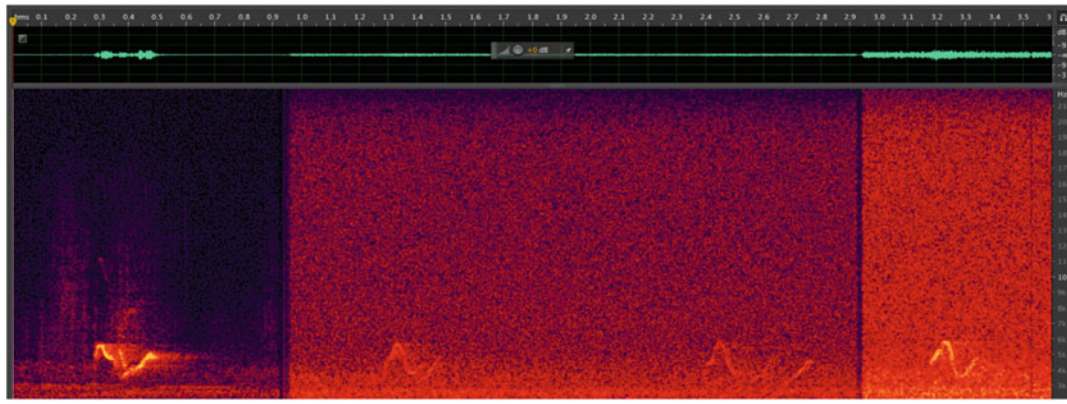


**Fig. 2.3** Spectrogram depicting high self-noise versus low self-noise output by three microphone/recorder combinations. In the left section, a low-noise system was used and the signal clearly emerged from the environment background. In the following sections, nosier systems were used; the sounds appear unclear and listening was unpleasant

an animal's sound that is outside the recording system's sensitivity or frequency range. This might especially happen, if the sound is above or below the human hearing range. For example, elephants communicate with conspecifics using infrasounds (Payne et al. 1986), and rodents and bats produce ultrasounds for communication and foraging (see Chap. 12 on echolocation).

Other features to consider when purchasing equipment for fieldwork are the construction quality, weather proofing, reliability, visibility of the display, and ease of use in harsh conditions (see Chap. 3 on practical considerations). Powering the instruments might be a major issue with regard to practicality, cost, and safety. For example, low-noise preamplifiers generally require higher operating currents. Large-capacity batteries increase the risk of fire. During long field trips, internal rechargeable batteries may be difficult to recharge; replaceable batteries may be easier to manage, and external powering options could become a necessity (e.g., to power a recorder with a standard 5 V USB source or with a 6- or 12-V battery pack). For extended autonomous deployments, the cost of the power source might end up exceeding the cost of the recording equipment.

### 2.3.1 Sensors

Microphones and hydrophones convert sound pressure signals into electrical signals. The electrical signal, which is representative of the original sound waveform, can be amplified, filtered, recorded, visualized, and further analyzed or converted back to sound for playback or projection. Speakers work in the reverse and convert the electrical signal into sound for broadcast. A transducer converts a signal from one form (of energy) to another. So microphones, hydrophones, and speakers are all transducers. Usually, microphones and hydrophones, as long as they do not have a built-in preamplifier, can be used as both sound sensors and sound projectors. But their receiving and projecting amplitude sensitivities, frequency responses, and directionalities may differ.

Each microphone and hydrophone has a unique amplitude sensitivity, frequency response, and directivity pattern. These are specified in the specification sheets of high-quality sound sensors. A flat frequency response gives the least distorted audio-signal; however, during signal calibration, a non-flat response can be accounted for. The sensor size influences amplitude sensitivity, frequency response, and

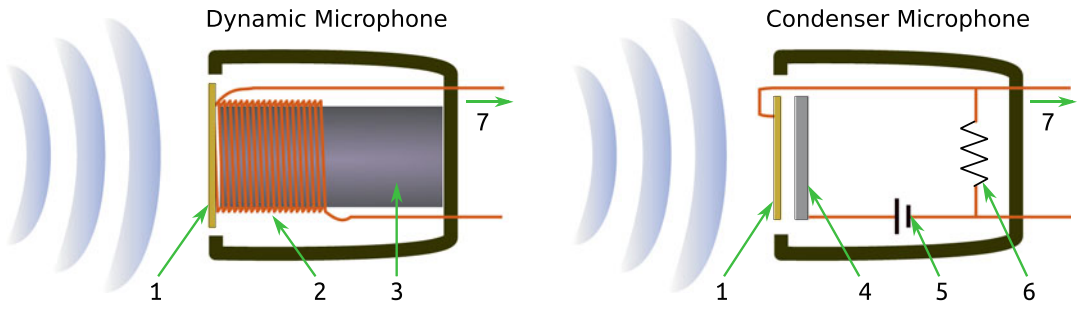



**Fig. 2.4** Schematic of a dynamic microphone (left) and a condenser microphone (right) showing the conversion of sound waves into electrical audio-signal outputs. Microphone schematic components: 1. vibrating diaphragm, 2. coil attached to the diaphragm, 3. magnet, 4. backplate, 5. battery, 6. resistor, 7. output

directionality. A sound sensor, to be omnidirectional, should be smaller than the minimum wavelength of the signal to be received. Large sensors are more sensitive but tend to limit responses at high frequencies. Large sensors become directional at lower frequencies than small sensors do.

#### 2.3.1.1 Microphones

Microphones convert sound energy (from sound waves) into an electrical audio-signal using a moving diaphragm or membrane. Two main types of microphones are common: dynamic microphones and electrostatic microphones (condenser and electret microphones) (Brüel and Kjær 1982). Some microphones are sensitive to particle motion, as well as sound pressure, which results in them being very sensitive to sounds very close to the microphone (i.e., in the near-field). This often exaggerates the low-frequency components of the received sound.

In dynamic microphones, a coil on the back of the diaphragm is immersed in a magnetic field and generates a current by electromagnetic induction when the membrane moves (Fig. 2.4). Such microphones do not require external power, but they have limited sensitivity, making them most useful for loud signals or at close range to the sound source. The delicate mechanical suspension in dynamic microphones may warrant gentle handling.

Electrostatic microphones are based on a condenser with a thin moving diaphragm (Fig. 2.4). Movement of the diaphragm changes capacitance in the condenser. Capacitance changes are then converted to voltage. Condenser microphones need a high voltage to polarize the condenser. In contrast, electret microphones are permanently polarized as their diaphragms are made of metallic-coated, pre-polarized, plastic membrane. Both condenser and electret microphones need power for their integrated preamplifier, with condenser microphones requiring additional power to polarize the condenser. This power may be supplied by an internal 3–5 V battery, 48-V phantom power (P48), or a Power-In-Plug (PIP) unit. P48 is a standard means of feeding power to a condenser microphone with 48 Vdc and is commonly used in professional recorders. Modern pocket digital recorders use PIP units for powering their microphones. The membranes in electrostatic microphones are delicate and sensitive to humidity, which can be problematic in humid environments. The lower mass of electrostatic elements generally yields superior high-frequency response. However, electrostatic sensors may be noisier than dynamic sensors. For studies involving low-frequency sounds, dynamic sensors may be a better choice.

A radio-frequency microphone is a special type of condenser microphone, developed by Sennheiser[2] in its MKH series. With this type of microphone, variations of the capacitor modulate the frequency of a radio-frequency oscillator, and then a demodulator extracts the audio-signal to be

[2] http://www.sennheiser.com/; accessed 15 Mar. 2021.

transmitted over a cable. The radio-frequency oscillator and the demodulator are both housed inside the microphone, and these microphones are less prone to problems of interference and humidity.

The more recently developed Micro-Electrical-Mechanical System (MEMS) microphones have pressure-sensitive elements integrated directly into a silicon chip (as found in most cell phones) with similar fabrication technologies used to make semi-conductor devices. Some integrate an AD-converter to produce a digital output. Their development resulted from the need for tiny microphones for cell phones. Because of the small size and low inertia of their sensors, MEMS microphones are sensitive to high frequencies and consequently are used in ultrasonic microphones, such as in bat detectors. Because of their low cost, they are the perfect candidates for array applications, including "acoustic cameras" that overlay the image taken by a video-camera with a map of the sound sources generated by a matrix of tens or hundreds of MEMS microphones.

Most condenser microphones have a self-noise lower than 20 dB(A), which is sufficient to record music or speech at a close distance, but not suited to record faint animal sounds and noises in a quiet environment. The quietest studio microphones have a self-noise below 10 dB(A); among these microphones is the Rode NT1A, a cardioid microphone that has an excellent self-noise of only 5.5 dB(A). Even quieter microphones are available in the category of instrumentation microphones, but few very expensive models are available. Lynch et al. (2011) and Pavan (2017) used very quiet instruments to show that noise in natural environments can be as low as 10 dB re 20 μPa and even go below 0 dB re 20 μPa below 1 kHz. Of course, a quiet microphone must be connected to a quiet recorder!

Sometimes, microphone specifications are difficult to read or self-noise is not provided. One must examine the parameters that are given, such as amplitude sensitivity and the signal-to-noise ratio (SNR). If not differently declared, the SNR is relative to 94 dB re 20 μPa (i.e., 1 Pa) at 1 kHz and thus the self-noise can be obtained by subtracting the given SNR from 94. If properly measured and reported, an SNR of 80 dB (A) means a self-noise of 14 dB(A), which is pretty good. In other cases, the sensitivity, the maximum allowed SPL, and the dynamic range are presented. In this case, the self-noise can be obtained by subtracting the dynamic range from the maximum allowed SPL.

### Ultrasonic and Infrasonic Microphones

Microphones for ultrasounds are typically small, with a small membrane with very low inertia. Ultrasonic microphones are usually condenser microphones developed for measurement purposes, not for recording music; however, the increasing interest in ultrasonic communication and echolocation in animals (mainly bats and rodents, but also insects) has fostered the development of a wide range of sensors for ultrasounds. Ultrasonic microphones for measurement purpose need to have a flat frequency response; usually they also have high self-noise and are very expensive. If the flatness of the frequency response is not a necessity, other, lower-cost microphones can be used instead (e.g., low-cost small condenser microphones and tiny MEMS microphones). Considering that ultrasonic microphones need high sampling rates, often beyond those available in consumer digital recorders or AD-converters (see Sect. 2.3.4), ultrasonic sensors with integrated AD-converter and USB interface have been developed. In bioacoustic studies, these are mainly used for detecting and recording bats (Sect. 2.3.5), insects (Buzzetti et al. 2020), and rodents either in the wild or in ethopharmacological studies (Buck et al. 2014).

Infrasonic microphones are specially designed for low-frequency recording, down to 1 Hz or even 0.1 Hz. Until a few decades ago, Sennheiser produced the MKH 110, a condenser microphone with 12-V powering. Now discontinued, it is still appreciated in the used equipment market. These microphones have been widely used to record elephant communication (Payne et al. 1986; Poole et al. 1988). Currently, microphones designed for infrasonic applications are largely

limited to measurement (instrumentation) microphones.

### Measurement and Specialty Microphones

Measurement microphones (or, instrumentation microphones) are a special class of microphones designed to make accurate measurements of sound amplitude within a specified frequency range, which could be infrasound to ultrasound, to accurately characterize a sound field or a sound source. These microphones comply with specific and rigid requirements. They need to have a well-defined and stable frequency response to sound (ideally flat). They usually appear as cylinders with diameters ranging from 1/8 inch for very high frequencies (but with low sensitivity) to 2 inches for high sensitivity and low noise (but limited extension to high frequencies). Normally based on condenser sensors, these microphones are often powered at 200 V. Measurement microphones are usually connected to specific digital recorders and analyzers, or integrated into a sound level meter (also known as phonometer). Usually dedicated to noise measurement, these microphones are also used to calibrate other types of instruments (see Sect. 2.6) and to record sounds for analysis and listening with great accuracy. Brüel & Kjær[3] are well known for their measurement microphones; however, other manufacturers exist as well, providing a wide range of sensors for applications of sound recording, acoustic measurements, noise monitoring, building acoustics, cinema calibration, occupational health, and live sound broadcasts.

Optical microphones are a very special category of measurement microphones. A laser beam is reflected by a very tiny low-inertia sound-sensing membrane, and the reflected beam is then detected by an optical sensor to extract the modulation given by the membrane moved by sound waves. Their advantage is the direct optical output that is conducive for long-range transmission over optical cables and their insensitivity to electric and electromagnetic fields.

[3] http://www.bksv.com/en/; accessed 15 Mar. 2021.

Wireless microphones transmit the received sound by a radio signal that can be either a standard AM- or FM-transmission or a digital format to ensure signal quality and privacy. Wireless microphones allow the cable-less transmission in situations where cables are problematic. Wireless microphones connected to a multi-channel receiver allow a wide area to be monitored. In some cases, the wireless microphones used for television interviews can be used successfully (e.g., by placing the microphone close to or inside a nest and then recording from a distance). A traditional microphone can also be equipped with a radio transmitter and a battery that powers both. The limitations include powering the transmitters (in particular, in field and long-term deployments), limited dynamic range, compromised self-noise, and radio-frequency interference during transmission.

### Microphone Directionality

Directionality is an important characteristic of a microphone. Omnidirectional microphones detect sound from all directions and can be appropriately used for recording a soundscape (i.e., the combination of all sounds generated in an environment; see Chap. 7). Directional microphones are good for making recordings of a selected animal in a specific direction (e.g., a particular individual in a colony) and for attenuating noise coming from directions other than the signal direction (e.g., the noise of a nearby river or road). Directional microphones thus improve the SNR by reducing background sounds and noise coming from other directions in the environment. In indoor applications, directional microphones are used to focus on a performer and to attenuate reverberation from the hall. Widely available types of directional microphones include cardioid, hypercardioid, bidirectional, and unidirectional (Fig. 2.5). Cardioid microphones exhibit a heart-shaped directivity (i.e., they are less sensitive at 180° from the sound source) and they are often used with parabolic reflectors. The hypercardioid microphone is less sensitive at ±120° from the direction to the sound source. Bidirectional microphones pick up sound in a

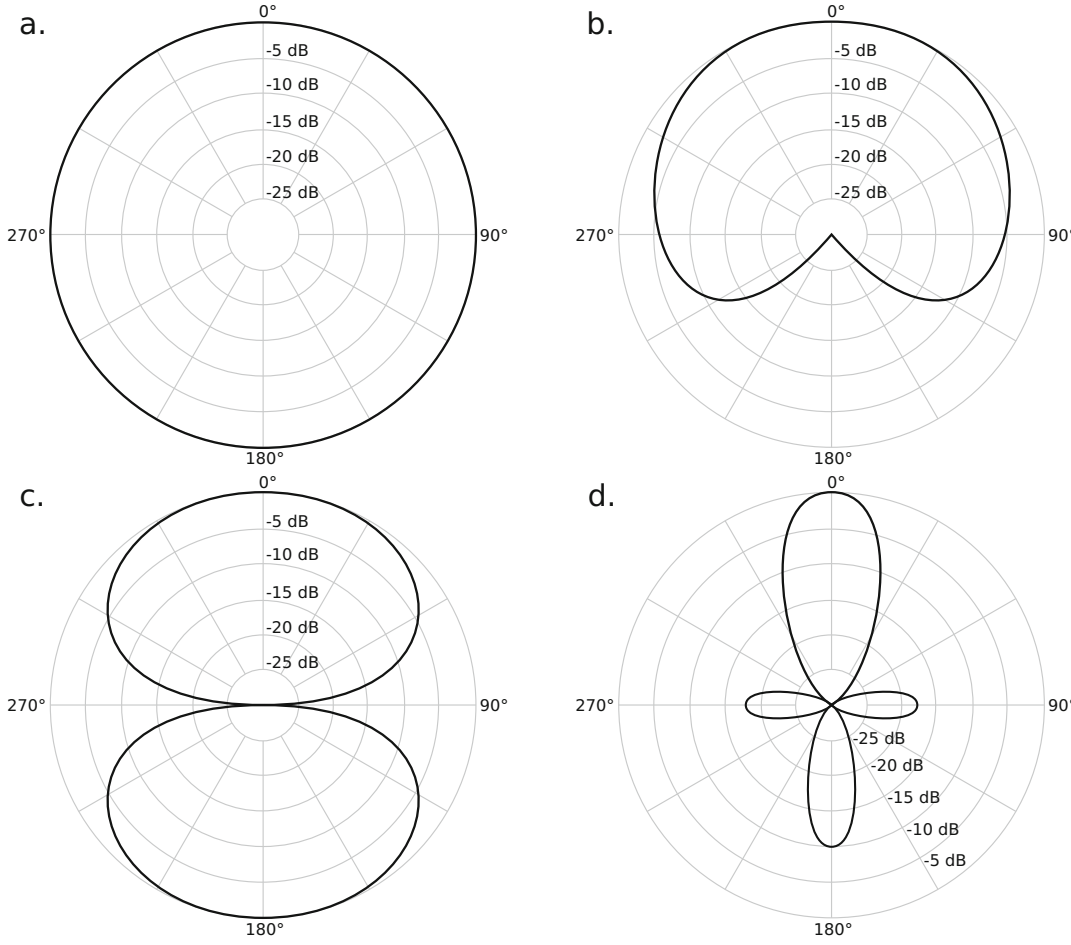



**Fig. 2.5** Polar patterns of directionality of different microphones. With microphones facing the top of the page, these patterns extend from the axis of the microphones, and thus present directivity in the vertical plane. In the horizontal plane, these patterns are symmetrical (i.e., they rotate about the vertical axis). (**a**) omnidirectional, (**b**) cardioid, (**c**) bidirectional (figure-of-8), and (**d**) shotgun (lobar)

figure-of-8 pattern equally from two, opposite directions.

Shotgun microphones (Fig. 2.5d) are the most directional and commonly used for recording a specific animal. Their use is desirable when it is necessary to improve the recording level of a specific sound source, or to attenuate unwanted sound coming from other directions. The design of shotgun microphones (such as the Sennheiser K6/ME66 or the MKH 8070) is based on the interference tube principle; usually a cardioid condenser microphone is placed at the end of a tube with slits on sides, canceling off-axis signals (Fig. 2.6). The directivity increases with the length of the interference tube and with the frequency of incoming signals, so that at high frequency ($>$ 4 kHz), the receiving lobe is quite narrow. For lower frequencies, the directivity decreases. This also means that off-axis sounds are not only attenuated, but also have a modified frequency spectrum, with high frequencies more attenuated than low frequencies. At wavelengths longer than tube length, off-axis attenuation is null. If interested in higher frequencies, such as bird songs above 1 kHz, a high-pass filter to cut off low frequencies (e.g., to attenuate wind noise or traffic noise below 150 Hz) is available in high-quality microphones.

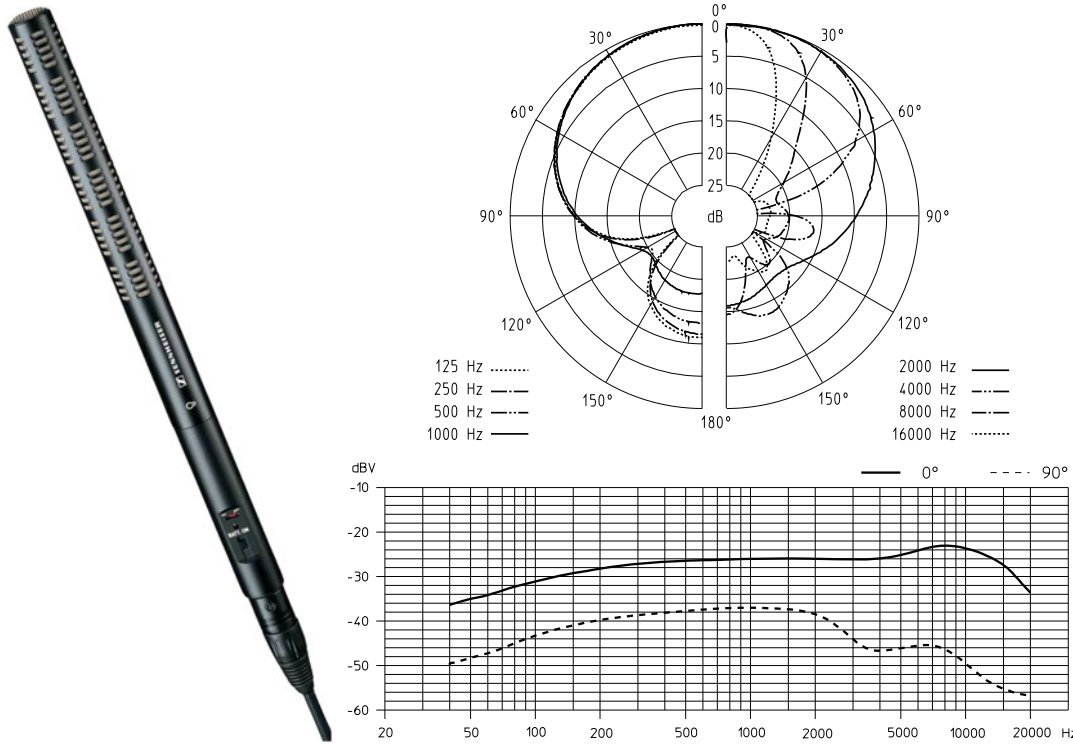



**Fig. 2.6** Photograph (left) of a modular microphone (Sennheiser K6/ME66) with the preamplifier body that hosts a battery to power the microphone in case the P48 powering is not available; the sensing capsule is interchangeable (omni ME62, cardioid ME64, short shotgun ME66, shotgun ME67). Polar pattern (top-right) of the microphone at different frequencies and the frequency response (bottom-right) on axis and at 90° from the sound. Reprinted with permission from Sennheiser

### Monophonic and Stereophonic Recording

Monaural recordings are made with a single microphone. Stereo recordings are made with two microphones and provide a sense of depth or movement through space in recordings. Stereo recording offers spatial information, which helps better discriminate sound sources in the surrounding space. Three primary setups are used for stereo recordings (Fig. 2.7): XY, binaural, and MS (middle-side). A common setup for the XY stereo recording uses two cardioid or super-cardioid microphones placed at 60° or 90° angles, nose-to-nose. The two microphones can be coincident or spaced. In some cases, the left microphone points in the left direction, in other cases, the left microphone points in the right direction and the right one in the left direction.

In the binaural stereo recording configuration, two omnidirectional microphones are placed approximately the distance between the ears of a typical human head (16–18 cm spacing) through the use of a mannequin head that simulates a human head and ears. This presents the idea of three-dimensional (3D) sound experience as the listeners with headphones have the sensation "to be there," with their ears in the same position of the microphones. The microphones can also be separated with nothing in-between, or with just a generic separation, such as a sphere of foam, or a Jecklin disk. Another special binaural configuration is called the Stereo Ambient Sampling System (SASS) design that simulates a human head. Compared with other techniques, with exception of the true binaural, this type of recording

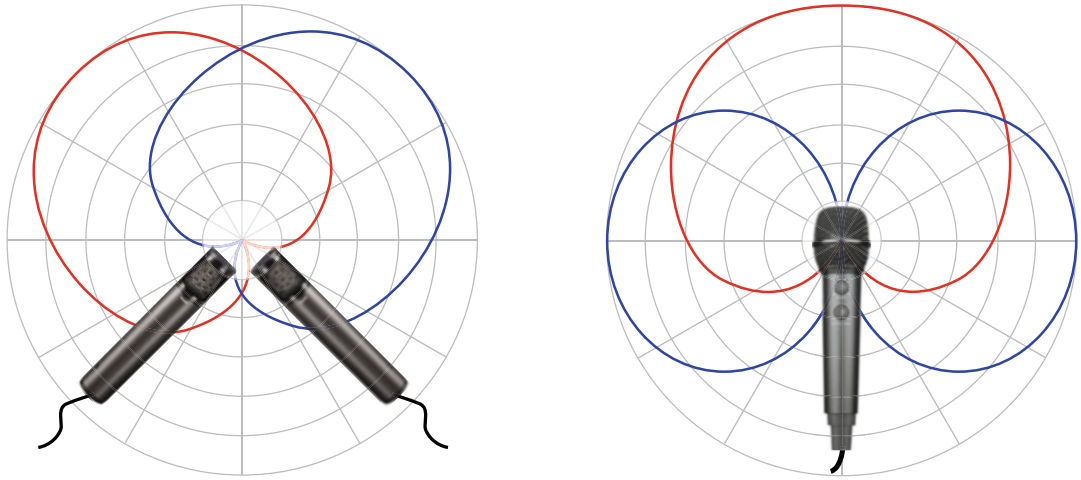


**Fig. 2.7** XY recording configuration (left) using two cardioid microphones, and MS recording configuration (right) which typically combines a cardioid microphone in the middle and a bidirectional microphone taking the sounds coming from the sides (figure-of-8 polar pattern)

produces the best spatial image when heard through headphones. In some setups, cardioid microphones angled at 60°–90°, like in the XY configuration, are used to enhance left-right separation.

In the MS microphone stereo recording setup, a cardioid microphone is piggy-backed on top of a bidirectional microphone. The cardioid picks up frontal information, whereas the bidirectional microphone gets sounds coming from the sides only. This type of recording requires specific electronics, or signal processing to combine the signals to produce a traditional stereo image. In essence, the signals from the left and right capsules are summed out-of-phase before being combined with the mono-signal. This computation allows the recordist to control the width of the stereo spread and make other adjustments in post-processing. In the early stages of the sound industry, this helped to maintain the compatibility among mono and stereo recordings. Several microphone arrangements have been developed for stereophonic recording; for a comprehensive review, see Rayburn (2011) or Streicher and Everest (1998).

Latest developments, mainly driven by the film industry to produce an immersive 3D (full-sphere, surround-sound) acoustic environment, capture sound not only in the horizontal plane, but also above and below the listener. Surround-sound recording requires several microphones in a 3D configuration, whose signals (channels) are electronically or digitally combined to produce both stereo and multi-channel surround-sound experiences, or to create specific receiving beams (e.g., to focus on a sub-space or on a specific source). The Ambisonics system allows recording of sound pressure on 3 axes with 4 microphone capsules mounted as a small tetrahedron (first order Ambisonics) (Zotter and Frank 2019). Higher-order Ambisonics microphones can have up to 32 capsules on a small sphere to achieve higher directional details and to simulate virtual directional microphones to be oriented in any direction during post-processing.

### Microphone Arrays

Arrays of sound sensors are used to monitor animals across habitats, locate and track sound sources (such as individual animals), and study environmental noise. Arrays may be stationary (fixed in location), freely drifting (e.g., suspended from balloons), or towed. Ambisonic microphones, are a special case of microphone

arrays. The sensors in an array operate in tandem. Their signals are combined in digital signal processing. A number of requirements need to be met for successful array processing (e.g., to track a bat by its biosonar). Sensor locations need to be known accurately. Sensor directionality needs to be known. Sensor spacing must be such that the target signal can be detected on multiple sensors. These sensors need to be matched and their eccentricities need to be computed. Time differences of arrival (TDOA) need to be computed between sensors. An overview of digital signal processing algorithms to locate and track sound sources is given in Chap. 4.

While the complexity of meeting the above requirements has limited the application of microphone arrays for animal localization and tracking in terrestrial environments, Mennill et al. (2012) successfully deployed an array of wireless microphones with integrated Global Positioning System (GPS) time synchronization to make accurate measurements of the position of a sound source by computing TDOAs of the same sound at different microphones. They discuss how this system may be implemented to monitor frogs, birds, and mammals. Jensen and Miller (1999) used a 13.5-m vertical, linear microphone array that allowed for simultaneous recordings of bat signals at three different heights of vegetation. With this design, they were able to calculate flight direction, altitude, and distance from the array.

The literature sometimes presents arrays of sensors that do not operate in tandem. Rather, sensors are widely spaced over a potentially large area, sampling independently without synchronization. The applications are not to locate and track individual sound sources, but rather to monitor a soundscape, compare animal presence/absence across sites, or evaluate environmental noise over a large area. During digital signal processing, noise levels might be compared across sites and perhaps interpolated to produce a noise map. For example, the Cornell Lab of Ornithology uses an array of 30 recorders to monitor animal habitat use on a wide spatial scale and to assess anthropogenic impacts (Fig. 2.8).

### Do-it-Yourself (DIY) Microphones

Microphones well-suited for bioacoustic studies can be built with microphone capsules costing only a few US dollars. Examples are the omnidirectional electret capsules from Primo Microphones Inc. (EM models)[4] or the PUI Audio Inc. AOM-5024 L model.[5] These capsules can be powered directly by PIP when connected to a handheld digital recorder, or powered with a battery and a simple electronic circuit. Adapters can be easily built to power PIP microphones with the P48 powering provided by professional recorders that do not provide PIP.[6] DIY microphones can be easily assembled to experiment with different spatial configurations, even in the focus of a parabolic reflector, or to have low-cost expendable microphones for very specific field tasks.

### Deployment Considerations

In open-field environments, wind can affect signal reception by a microphone by causing non-acoustic noise, which is an artifact of turbulent pressure fluctuations at the external surface of the microphone. Such turbulent pressure fluctuations may be caused by the obstruction that the microphone itself presents. Turbulent air flow may also be caused elsewhere and produce noise artifacts in recordings as the perturbations travel past the microphone. Even a light breeze can produce strong low-frequency noise artifacts, which can overload the internal electronics or the recorder. Microphones can be fitted with a windsock to reduce wind noise. A windsock can be easily made with commercially available open-cell foam, which limits air flow but allows sound waves to reach the microphone membrane. For severe wind conditions, a fur-like cover is preferable (Fig. 2.9).

When aiming to record animals in a specific direction (e.g., a bird calling from a tree), a directional microphone should be used and pointed at

[4] https://www.primomic.com/; accessed 15 Mar. 2021.

[5] https://www.puiaudio.com/; accessed 13 Aug. 2021.

[6] http://tombenedict.wordpress.com/2016/03/05/diy-microphone-em172-capsule-and-xlr-plug/; accessed 13 Aug. 2021.

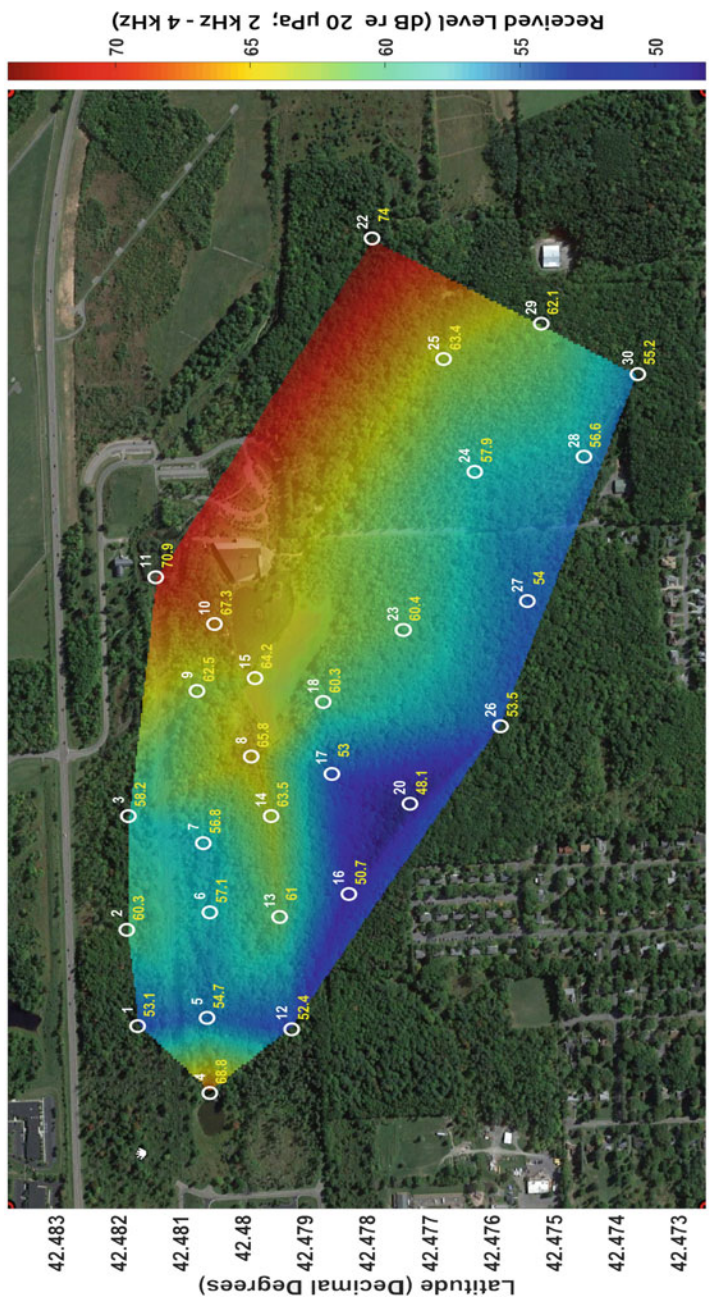



**Fig. 2.8** Noise pattern observed at Sapsucker Woods (Ithaca, NY, USA), caused by a jet plane taking off from a nearby runway. Receiver locations are denoted by white circles. Regions shaded red show high noise levels and follow parallel to the flight path. The ambient noise levels are raised by about 15 dB at the frequencies of chorusing birds (2 to 4 kHz). Image courtesy of Dimitri Ponirakis, K. Lisa Yang Center for Conservation Bioacoustics, Cornell Lab of Ornithology



**Fig. 2.9** Photograph of a microphone setup with pistol grip and elastic suspension, foam windsock, and additional furry windsock for maximum wind protection. Reprinted with permission from Sennheiser

the bird. It will focus sound recording in the direction of the bird and limit background noise from other directions. An alternative to a highly directional shotgun microphone is a cardioid microphone placed in the focus of a parabolic reflector (Fig. 2.10). The microphone is pointed toward the parabolic reflector, facing into the dish, not toward the animal. Ideally, the microphone's beam pattern would be matched to the solid angle subtended by the reflector. The diameter of the parabolic reflector determines which frequency range of incoming sounds will be amplified (Fig. 2.11). To be reflected, the wavelength of the incoming sound must fit inside the dish. The lowest frequency a parabola can reflect, and thus focus on the microphone, depends on the dish diameter (Wahlstrom 1985). For a 1-kHz signal, a 30.5 cm diameter dish is fine, and for a 500-Hz signal, a dish of 61 cm in diameter is required. The very low frequency of a lion roar (40–200 Hz) would require a dish about 10 m in diameter.

Compared to shotgun microphones, parabolic reflectors intercept a much wider quantity (proportional to the diameter and surface of the reflector) of acoustic energy and concentrate it on the microphone, thus providing a high gain. However, this gain is proportional to the frequency and the parabola diameter, thus producing a recording with increased high-frequency levels that requires equalization in post-processing (some parabolas can have equalization built-in). As a rule of thumb, the more wavelengths are contained in the parabola diameter, the higher the gain and greater the directionality. Because of these features, parabolas, with the right choice of microphones, can provide excellent recordings of very quiet, distant sources. For example, in a taxonomic and behavioral study of chipmunks (*Neotamias* spp.), Gannon and Lawlor (1989) used a 51-cm parabolic reflector with a Sennheiser ME-20 omnidirectional microphone and K3U preamplifier. Chipmunk calls were in the range of 4 kHz to 15 kHz, so this size dish was

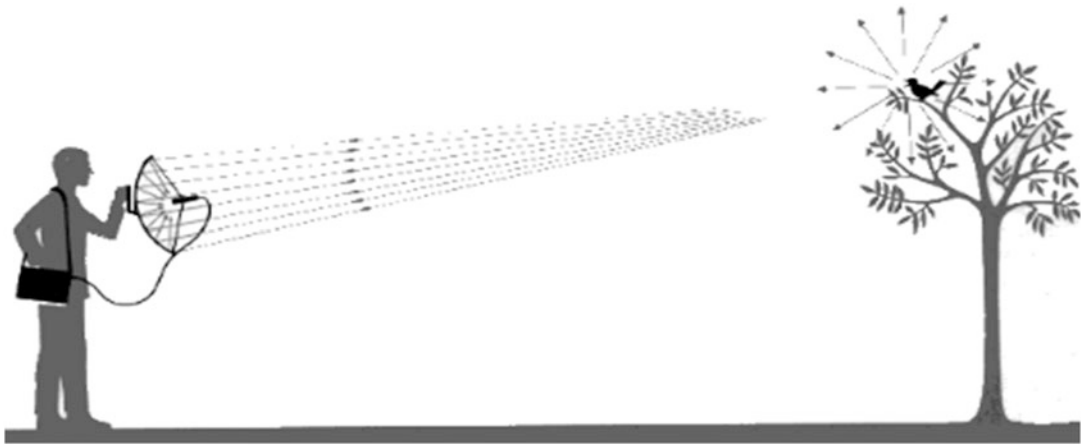


**Fig. 2.10** Diagram of a parabolic dish and microphone used to record a bird on a tree. The parabolic solution gives added amplification and directivity, which helps in recording a single animal, a quiet animal, or animals at a distance

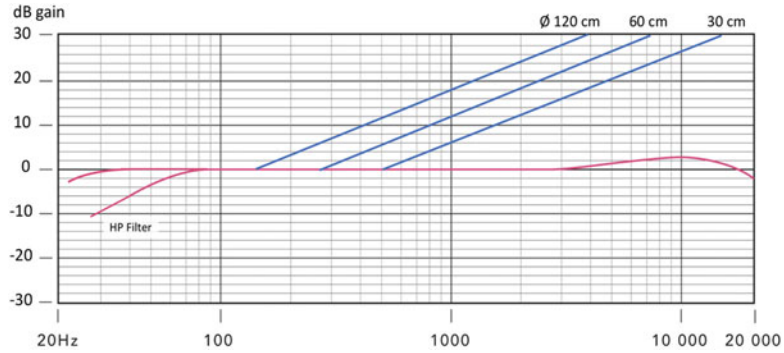



**Fig. 2.11** Sketch of frequency response and gain of a generic microphone placed in parabolas of different diameters. The red lines show the frequency response of an ideal microphone, with the option of a high-pass filter to reduce low-frequency noise below 80 Hz. The blue lines show the theoretical gain of three parabolas of different sizes. The gain is proportional to frequency and to the parabola diameter. Actual response may vary depending on the shape and depth of the parabola and on the response and positioning of the microphone

adequate for detecting this range of mid-frequency calls.

To produce a more pleasant recording, it is possible to record in stereo by using two microphones in the focus, separated by a thin plate. This way, sounds coming from the frontal axis of the parabola reach both microphones with the same level, while off-axis sounds are focused more on one side. Another option is to place an MS microphone combination in the focus of the parabola. Listening with headphones helps in pointing the parabola on the source of interest and gives immediate feedback on the quality of the sounds being recorded. When analyzing recordings made with a parabola, it is important to take into account that the frequency response is not flat as it increases with frequency (Fig. 2.11). In some cases, slightly moving the microphone out of focus reduces the high-frequency emphasis and produces a more pleasant sound.

#### 2.3.1.2 Hydrophones

A hydrophone is a piezoelectric transducer that converts sound waves in water to electrical signals. Hydrophones can receive sound in air, but the sound has to be of very high amplitude. Because the acoustic impedances of the medium and the sensor match much better in water than in air, hydrophones have to be less sensitive, or they would easily overload. The underwater sensor usually is sealed in a resin package with a waterproof connector and needs to be handled with care. After use in saltwater, a hydrophone should be rinsed with freshwater or else connections are likely to corrode.

A piezoelectric transducer can be used as a sensor or projector; however, when the transducer has a built-in preamplifier, it can no longer be used as a projector, but only as a sensor. Hydrophones are much less sensitive, and a great deal of power is needed (from an external amplifier) to drive a hydrophone as a projector. As a sensor, a hydrophone can have a built-in preamplifier that matches the frequency response, dynamic range, and high impedance of the transducer. A few hydrophones on the market with built-in preamplifier (Fig. 2.12) can be powered directly by a recorder, computer, or analysis system (e.g., either by P48 or by PIP at 2–5 Vdc). Most preamplified hydrophones require powering through dedicated cables and can require single or dual powering (e.g., +12 V, or −12 V and +12 V) to be provided by a battery box (Fig. 2.12). A popular low-cost hydrophone is the H2c from Aquarian Audio,[7] which allows PIP powering. The DolphinEar[8] is an inexpensive, lightweight,

[7] http://www.aquarianaudio.com/; accessed 15 Mar. 2021.
[8] http://www.dolphinearglobal.com/; accessed 19 Jun. 2022.



**Fig. 2.12** Photographs of an ITC 6050C hydrophone with built-in preamplifier and external battery power (left) and a Cetacean Research Technology C57 hydrophone with cable and battery box (right; courtesy of J R Olson)

battery-operated hydrophone with an external amplifier and headset that is good for ecotourism or classroom use. Other relatively low-cost hydrophones well suited for marine mammal studies are produced by Cetacean Research Technology.[9]

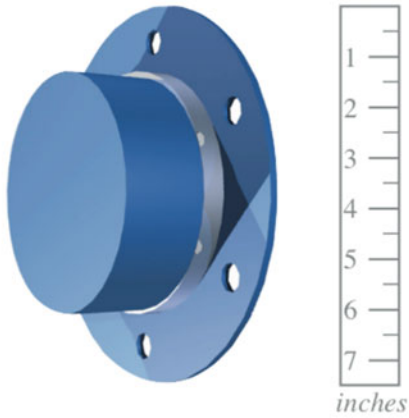
To record underwater sound in open water from a distant source, a sensitive hydrophone is needed. Good sensitivity would be −160 dB re 1 V/μPa. Such a hydrophone produces 1 V when receiving 160 dB re 1 μPa of acoustic pressure and 1 mV for a signal of 100 dB re 1 μPa. If used for recording a signal at 180 dB re 1 μPa, it will produce a 10-V output and may overload the connected electronics. To record underwater sound at close distance (e.g., in front of an echolocating dolphin which can produce pulses with source levels above 220 dB re 1 μPa m pk-pk), a low-sensitivity hydrophone is needed (e.g., one that has a sensitivity of −210 dB re 1 V/μPa). Very likely, such a hydrophone cannot be used for recording low-level sounds from a distant source because it requires high amplification and consequently produces high electronic noise. However, using hydrophones with built-in preamplifiers when powerful signals can occur risks overloading of the preamplifier, thus producing distorted signals. Erbe (2009) used four different hydrophone systems (differing in amplitude sensitivity) to record impulsive pile driving at ranges from 14 m to 1330 m.

Hydrophones can vary considerably in their frequency response; some are used specifically for low-frequency, mid-frequency, or high-frequency reception. Typically, hydrophones are smaller than the wavelengths that are being recorded. But, with the smaller sensor comes a lower energy input. This results in lowered sensitivity. Generally, the smaller the piezoelectric element, the broader the frequency range, but the lower the amplitude sensitivity. Lower sensitivity can require higher amplification, and thus can produce higher electronic noise. Piezoelectric hydrophones usually have a resonance peak in the upper part of their bandwidth, so that optimum operation of the hydrophone is along the flat portion of the frequency response curve below resonance. Reception at other frequencies could be used, but the difference in response of the hydrophone needs to be accounted for during analyses. Some studies require the use of multiple hydrophones to cover the entire frequency range of the animal's sounds.

### Hydrophone Directionality

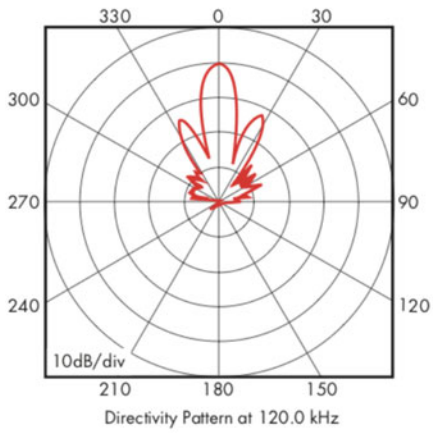
Hydrophones, much like microphones, have directional receiving and transmitting characteristics, depending on the size and shape of the transducer (Fig. 2.13). Spherical transducers receive and transmit signals uniformly in all directions. With a cylindrical

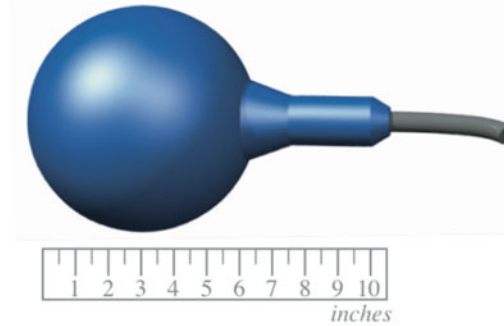
[9] http://www.cetaceanresearch.com/; accessed 15 Mar. 2021.




**Specifications (Nominal)**

| | |
|---|---|
| **Resonance Frequency $f_r$** | 118 kHz |
| **Depth** | Unlimited |
| **Envelope Dimensions (in.)** | 2.25D x 7.5H |
| **TVR at $f_r$** | 160 dB/µPa/v@1m |
| **Beam Width (-3dB) at $f_r$** | 10 deg |
| **Beam Type** | Conical |
| **Input Power** | 1000 watts |




**Specifications (Nominal)**

| | |
|---|---|
| **Type** | Projector/Hydrophone |
| **Resonance Frequency $f_r$** | 11.5 kHz |
| **Depth** | 1250 meters |
| **Envelope Dimensions (in.)** | 6.5D |
| **TVR at $f_r$** | 149 dB//µPa/V@1m |
| **Midband OCV** | -188 dB//1V/µPa |
| **Suggested Band** | .01 – 20 kHz |
| **Beam Type** | Spherical |
| **Input Power** | 10,000 watts |

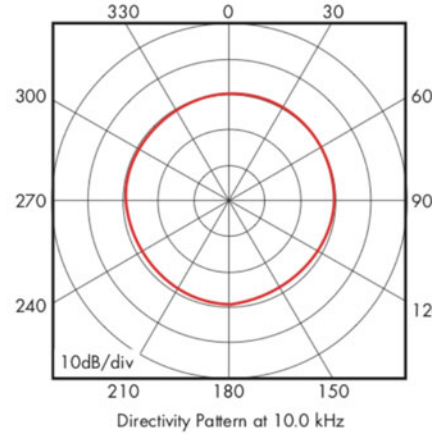



**Fig. 2.13** Specifications and polar plot of directional ITC 3003D transducer (left) and omnidirectional ITC 1007 transducer (right). Reprinted with permission from Gavial ITC (https://www.gavial.com/itc-products; accessed 22 Aug. 2021)

transducer, sounds are received and projected uniformly in the horizontal plane, assuming the transducer is suspended vertically. In the vertical plane, the transducer will have a directivity pattern. If the transducer has a planar shape, it will have two beams on its opposite faces as shown in the left polar plot in Fig. 2.13. When used as a sensor, a spherical hydrophone is typically omnidirectional (receives sounds equally from all directions) as shown by the right polar plot of Fig. 2.13. Used as a projector, the directivity pattern of a hydrophone changes depending on the frequency being projected (directivity increases with frequency).

### Sonobuoys

A sonobuoy is a canister housing a hydrophone, dampening cable, battery, recording/transmitting electronics, and a transmitting antenna. Navies of the world use sonobuoys for underwater listening by deploying them from aircraft or ships. These devices also may be used for bioacoustic studies. Once a sonobuoy is deployed in saltwater, a battery is activated, which triggers the inflation ($CO_2$) of a flotation balloon and antenna. The hydrophone and associated dampening cables can be set to drop to a pre-selected water depth (i.e., 30, 60, 120, or 300 m). During operation, the sonobuoy canister floats at the water surface with the antenna in the air and transmits acoustic data in real-time to a receiver onboard a vessel or aircraft or to a receiver at a station onshore. After a preset time (e.g., 1, 2, 4, or 8 h), a burn-wire penetrates the flotation balloon, and the sonobuoy fills with water and sinks to the seafloor.

Analog sonobuoys (Fig. 2.14) are available in two common configurations: omnidirectional sonobuoys (with a frequency response of up to 20 kHz) and DIrectional Frequency Analysis and Recording (DIFAR) sonobuoys, which provide bearing information on incoming signals. The latter type has been used to determine source levels and calling rates in cetaceans (e.g., Miller et al. 2015). The most recent generation of sonobuoys features a digital recording system and is equipped with GPS technology.

### Stationary Hydrophone Arrays

Stationary hydrophone array configurations include moorings (with or without surface buoy), seafloor packages, or cabled systems. Arrays of permanent, stationary hydrophones can be placed on the seafloor and connected via cables, either electrical or electro-optical, to processing centers located on shore. Multi-channel receivers allow listening or recording of sounds from multiple hydrophones. Typically, the array is optimized for long-range acoustic reception by using very-low-frequency sensors. Some bottom-mounted arrays are equipped with wideband hydrophones to allow scientists to monitor a wide variety of marine species, as well as ambient noise levels (e.g., Caruso et al. 2015; Favali et al. 2013; Nosengo 2009; Sciacca et al. 2015). Usually, these arrays are installed and maintained by navies, oceanographic organizations, or research centers for many years (see Chap. 1 for a list of past and current bottom-mounted hydrophone arrays deployed around the world).

### Towed Hydrophone Arrays

A towed array contains several hydrophones (not necessarily of the same type), commonly housed in an oil-filled sleeve (Fig. 2.15), where the oil matches the acoustic impedance of sea water. Originally developed for navies and geophysical survey companies, towed arrays were bulky and expensive, and mainly received low-frequency



**Fig. 2.14** Photograph of a sonobuoy deployed from a ship to monitor whale sounds in the Mediterranean Sea (SOLMAR Project, http://www.unipv.it/cibra/res_solmar_uk.html)

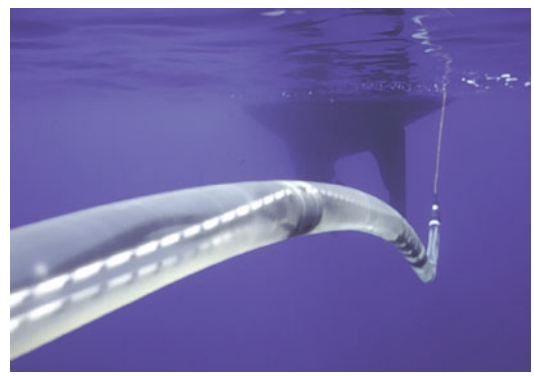


**Fig. 2.15** Photograph of a towed array under water, developed by the University of Pavia (Italy), with the tow vessel in the background

sound (<15 kHz). In more recent years, lightweight, wideband towed arrays sensitive up to 100 kHz and more have been developed to meet the requirements of researchers aiming to study marine mammals from small platforms, such as sailboats (Pavan and Borsani 1997; Pavan et al. 2013). By simultaneously processing sound from more than one hydrophone (or group of hydrophones), the bearing (or even location) of the vocalizing animal maybe be determined (see Chap. 4, section on sound localization). Towed arrays are used for line-transect surveys and to sample animals in their environment over a wide geographic range.

A straight-line array cannot resolve between signals arriving from the port or starboard side without the vessel changing course or using multiple array deployments (Thode et al. 2010). Large arrays (sometimes hundreds of sensors, possibly with different frequency sensitivities and bandwidths) allow tracking of multiple sources simultaneously by selective beamforming (Zimmer 2011). More complex towed systems use a 3D hydrophone configuration called a volumetric array (Zimmer 2013) or vector sensors (Thode et al. 2010) to locate sound sources in three dimensions. Acoustic vector sensors are sensitive to particle velocity rather than to pressure and hence sense the direction of incoming sound waves and resolve the directional ambiguities. Thode et al. (2010) attached a vector sensor module to the end of an 800-m towed array to detect sperm whale clicks and compute unambiguous bearing estimates of whales over time.

Many towed arrays have a depth sensor, so the operator knows the tow-depth in relation to the sound velocity profile in the water column. Such information allows the user to position the array either in a surface duct or below the thermocline to listen to sounds coming from deep water (see Chap. 6 on sound propagation under water). Additionally, the depth information enables subsequent array processing to exploit the surface effects on sound propagation to improve localization accuracy.

Array performance is degraded (in particular below ~1 kHz) by vessel self-noise, hydrodynamic noise artifacts (flow noise), and non-acoustic mechanical vibration, which reduce the ability to capture low-frequency animal sounds and which can cause an acoustic overload of the recording chain. To mitigate these issues, tow speed should usually not exceed 6 knots. A long cable with special elastic sections in the array can dampen vibrations. Flow- and vessel-noise can be mitigated with a smooth high-pass filter (e.g., 500 Hz, 12 dB/octave; see Sect. 2.3.2.1).

#### Deployment Considerations

To operate properly, hydrophones must have little vertical or horizontal movement. Water flow over the surface of the hydrophone generates pressure fluctuations, which appear as noise in spectrograms but which are not due to an acoustic wave. This flow noise is an artifact of deployment (see Chap. 3, section on flow noise). It is typically of low to mid frequencies (see, for example, the spectrogram in Fig. 3 in Erbe et al. (2015) showing flow noise in marine soundscape recordings) and thus can be filtered out with a high-pass filter, but this limits the recording of low-frequency sounds. Large or rapid vertical or horizontal movement of a hydrophone (e.g., if it is deployed over the side of a boat) may cause the system to be saturated with no useable recordings collected. It is very difficult to make good recordings in the open ocean; a hydrophone often needs to have its own flotation system, rather than be suspended from a boat; otherwise, the movement of the boat will translate into movement of the hydrophone. The horizontal component of water flow past a hydrophone may be minimized by deploying freely drifting hydrophone systems (e.g., suspended from a freely drifting buoy). The vertical component of water flow past a hydrophone may be minimized by dampening systems; for example, suspending the recorder on a bungee with a movement-dampening drogue, or by using a catenary floatation line (see Chap. 3 and Fig. 5 in Erbe et al. 2019). In towed arrays, long towing cables and specifically designed hydrophones (acceleration-compensated) are used to avoid saturation of the hydrophones from movement.

### 2.3.2 Filters

Filters are used to minimize unwanted noise from the environment (including other animals) or electronic self-noise. Filters can be used while recording or during post-processing. Filtering during recording facilitates conserving recorder dynamic range for signals in the frequency band of interest. A filter can be a stand-alone unit (some also have an amplifier) or filtering can be achieved using software, either in real-time or in post-processing. Note that filters are not a "magic wand" to make a bad recording clean. While recording, filters can be used to suppress unwanted noise without affecting the sounds of interest only when the noise and the sounds do not overlap in frequency. If noise and sounds do overlap (in frequency, or in time, or both), it is possible to perform some filtering or noise removal in post-processing. However, the settings need to be carefully chosen. Some microphones and digital recorders (Sect. 2.3.4) have built-in selectable filters, often with selectable attenuation rates.

#### 2.3.2.1 Low- and High-Pass Filters

Using a low-pass filter, the recordist can set a frequency above which signals are attenuated. A high-pass filter attenuates signals below a selected frequency. High-pass filters are often used to reduce low-frequency noise generated by wind and road traffic in terrestrial recordings and flow noise in underwater recordings. For example, to record a bird singing in the 2–5 kHz range, a high-pass filter set at 1 kHz will suppress traffic noise (which is typically below 500 Hz). A band-pass filter combines low-pass and high-pass filters. All filters have a transition bandwidth at the intersection of the pass band and the attenuation band, where there is a roll-off in the attenuation amount (steepness), which is normally expressed in dB/octave (e.g., 6 dB/octave in a smooth filter, or 24 dB/octave for a steeper filter). The greater the roll-off, the sharper the filter. However, sharper filters have longer impulse responses and generate longer artifacts in the output waveforms.

#### 2.3.2.2 Anti-Aliasing Filters

Digital recorders and audio interfaces have built-in anti-aliasing filters with varied performances; whereas instrumentation recorders and instrumentation acquisition boards usually do not have built-in anti-aliasing filters and require a separate signal-conditioning device to perform filtering and adjust the signal level. The available filters have their specific shape and thus can influence the frequency response of the recording.

AD-converters (Sect. 2.3.4) in recording equipment (either stand-alone recorders or external converters connected to a computer) have relatively smooth anti-aliasing filters that attenuate frequencies starting somewhat below the Nyquist frequency, but do not completely cut out the signal at Nyquist. Attenuation at Nyquist is often in the range of 6–12 dB, and the maximum attenuation (the FZero of the filter) is located above the Nyquist frequency.

The anti-aliasing filter shape is rarely reported in equipment specifications; tests are required to evaluate the anti-aliasing performances of the AD-converter, in particular if wideband signals are to be recorded and analyzed. Concern for aliased components is required for any type of signal possibly exceeding the Nyquist frequency, including external interferences captured by the electronics and cables, as well as higher harmonics of the signals to be recorded. A laboratory test with a frequency-generator signal sweeping across the whole frequency range of the recorder and beyond the Nyquist frequency can reveal unexpected and unwanted performance by the converter.

### 2.3.3 Amplifiers

A preamplifier conditions the incoming signal from a transducer and boosts the signal before it is recorded. A preamplifier converts a weak electrical signal into a stronger, noise-tolerant output signal for further processing. Without preamplification, the recorded signal could be noisy or distorted. The preamplifier has a high input-

impedance (i.e., it requires only a small current to sense the input signal) and a low output-impedance (so that when a current is drawn from the output, the change in the output voltage is minimal). In other words, a preamplifier converts a high-impedance input signal from a transducer to a low-impedance output signal. Besides lowering impedance, some preamplifiers also provide amplification (typically 20 to 26 dB). This is not true for most preamplifiers and hence they are typically paired with amplifiers. Preamplification should be constant across the recording bandwidth so as not to distort the signal. The frequency range and dynamic range specifications of the preamplifier and amplifier need to match other electronics in the recording system. For recording faint animal sounds or quiet soundscapes, the quality of the preamplifier is often an issue and must be considered carefully relative to the required use and the transducer to be connected.

An amplifier increases the signal gain after it is captured to drive the signal along a cable to the AD-converter without significantly degrading the SNR. Amplifiers can boost hydrophone signals as much as 60 dB (1000x). However, amplifying a signal will also increase ambient background sounds and self-noise; very high amplification could inadvertently make the noise level so high that desired signals cannot be recorded with good fidelity. Amplifiers for microphones are battery-powered and have high- and low-pass filters, which makes them useful for fieldwork.

Speakers include power amplifiers that drive a projector to generate high-amplitude acoustic signals in air or under water. The power amplifier provides the higher current to drive the speaker. Most power amplifiers used in high-fidelity home-entertainment systems also can be used in bioacoustic research. However, in some cases, more power and bandwidth are needed so that commercial broadcast power amplifiers must be used. No matter what class of amplifier or preamplifier is used, one should always consult the manufacturer's manual. Over-amplification can "blow" a loudspeaker or underwater projector.

### 2.3.4 Analog-to-Digital Converters and Digital Recorders

Despite declared sampling frequencies and bit-resolution, AD-converters, either in a stand-alone recorder or in a computer audio-interface, are based on diverse technologies and can affect the quality of a recording. For example, delta-sigma converters have high noise at high frequencies, beyond the human hearing limits, which becomes evident in wide-bandwidth power spectra and spectrograms. Another problem is jitter from instability of the clock driving the AD-converter and the digital stream. Excessive jitter can reduce the quality of recordings and can be seen easily by analyzing a clean test tone. Jitter can produce both random artifacts (Fig. 2.16) and periodic artifacts with well-defined frequencies. Jitter cannot be minimized by the user because it is characteristic of a given device. AD-converters can be divided into two main categories: for musical use, generally limited to the standard sampling frequencies of 44.1, 48, 96, and 192 kHz, or for instrumental measures, with sampling frequencies ranging from 100 Hz to 1 MHz and more. Converters for the consumer and prosumer musical market have smooth anti-aliasing filters included, suitable for musical signals, and a high-pass filter usually set below 10 Hz; instrumentation converters do not have any filter on their inputs and will sample any signal starting from 0 Hz (DC coupling). When using instrumentation converters, aliasing problems must be considered, and external anti-aliasing filters must be included in the recording chain (see Sect. 2.3.2.2).

An inexpensive and very portable AD-converter unit is PoScope's[10] Mega1 sampling at 500 kHz at 12 bit and recording directly to a PC in PCM files via USB interface. However, the PoScope, as most industrial data acquisition systems, including most National Instruments[11] devices, has no anti-aliasing filter and the measurement needs to be sampled at a rate much

[10] https://www.poscope.com/; accessed 15 Mar. 2021.

[11] http://www.ni.com/; accessed 22 Aug. 2021.

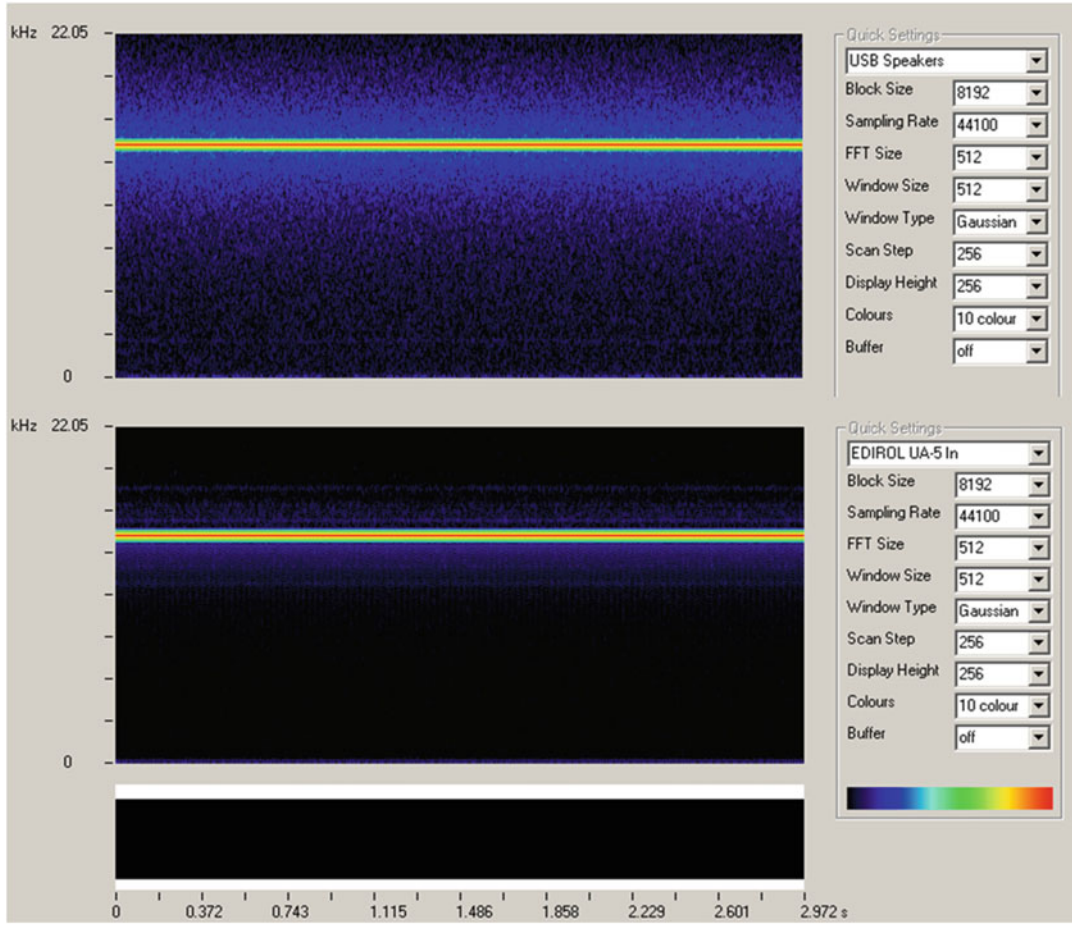



**Fig. 2.16** Spectrogram of a sinusoidal tone sampled at 44,100 Hz with a poor AD-converter (top panel). Note the low-intensity broadband noise (blue components) due to random jitter around the red line representing the tone's central frequency. Spectrogram of the same sinusoidal tone sampled at 44,100 Hz with a good AD-converter (middle panel); the broad blue band is absent in this image. The bottom panel shows the constant amplitude of the signal waveform

higher than the highest frequency contained in the input signals. If the upper-frequency content of the signal (including any possible noise or interference such as those generated by video monitors, digital networks, and switching power supplies) is unknown, use a good-quality, low-pass external filter at the known or presumed upper cut-off frequency while recording and digitally filter and down-sample the recorded file thereafter. It is also important to consider that strong low-frequency sounds below the desired frequency range can limit the dynamic range at higher frequencies of interest, so using a high-pass filter at a selected low frequency while recording is recommended.

AD-converters are more commonly available in the consumer market as "digital recorders" that also include the circuitry to save recorded data to permanent storage (e.g., SD-cards or internal memory) and an interface for powering the other components (either from an external source or through internal batteries). Some digital recorders also offer built-in selectable high-pass filters, which can help reduce the low-frequency noises

produced by handling and suppress wind or flow noises.

The frequency response of the digital recorder should be matched to the frequency response of the sensor–preamplifier–amplifier system as close as possible and to the needs of the research. The component with the narrowest frequency response is the limiting factor in the recording chain. All AD-converters have a maximum voltage range at the input that can be converted without overloading or clipping. The trick is to stay below the clip-level and still have good dynamic range and SNRs. Other important features in selecting the appropriate recorder are: the number of channels (e.g., 2, 4, 8, or more), durability, reliability for field-use, battery duration, flexibility and ease of use, maximum storage, integrated sensors (unidirectional or directional), inputs for external sensors, power options for the external sensors (P48 and/or PIP power), and the capability to connect a remote-control or a timer. Some recorders (especially many analog and digital tape recorders and video-cameras) use Automatic Gain Control (AGC) to keep the recorded volume within the same amplitude range. Other devices have an Auto Level Control (ALC) setting or a limiter function designed to avoid overloading or clipping. Some recorders indicate clipping either by a level-meter or with a flashing light. Any AGC, ALC, or limiter options should be disabled to perform comparisons among different sounds or different recordings and if true sound level measurements are needed. The gain level should remain constant throughout a recording, and noted; ideally, the sampling rate and gain settings should remain the same among recordings, at least for the same subject or context.

#### 2.3.4.1 Recording Ultrasounds and Infrasounds

Ultrasonic recorders were developed mainly for bat and dolphin studies; however, other animal species also produce ultrasonic sounds (e.g., insects, frogs, and infant rodents). To record ultrasound requires a sensor with suitable frequency extension and a recorder or an AD-converter with a high enough sampling frequency. An affordable solution is available in the form of ultrasonic microphones with integrated high-speed AD-converter and USB interface (e.g., Dodotronic[12] Ultramic family with sampling frequencies ranging from 200 kHz to 384 kHz). Dodotronic microphones do not need specific drivers and can be used on Windows, MacOS, and Linux, and also on Android smartphones. Recent models include support for internal storage (miniSD card) and powering with a USB battery box. The internal recorder can be set by Bluetooth to record on trigger or on a time schedule. Other similar devices are the Wildlife Acoustics Echo Meter Touch and Petterson Ultrasound Microphone. Another option for recording at very high sampling frequency is to use an instrumentation AD-converter like the PoScope Mega1+.

Many recorders are not suited for very-low-frequency recording. Most have a lower limit of 10–20 Hz; others can record down to 7–10 Hz. Recording very-low-frequency animal signals is complicated because this frequency range also contains environmental and electronic noise, which typically would be filtered out. For recording infrasounds (e.g., calls of elephants or baleen whales), it is important to check the specifications of the recorder and eventually make a bench-test of the available frequency range using a signal generator (a tone sweeping through a wide range of frequencies is a good test signal). An option is to use an instrumentation AD-converter with DC coupling.

#### 2.3.4.2 Special Features of Digital Recorders

Pre-recording buffer memory allows the user to save the few seconds of sound before pressing the record button. Auto-start initiates the recording automatically when a certain input level is exceeded. Double recording allows a lower-level backup copy in case some parts of the primary recording are overloaded. With this method, the incoming sound is recorded twice, in two different files, the second stereo file is stored at some dB down from the first file. In terrestrial

[12] http://www.dodotronic.com/; accessed 15 Mar. 2021.

applications, a wired remote-control can be useful when it is required to hide or protect the recorder (e.g., from rain). A wireless remote-control, by Bluetooth or by Wi-Fi (wireless fidelity), allows controlling the functions and levels by a smartphone application, but this would consume additional power and could impact energy budgets. File time-stamping inserts the date and time of the recording in the file name, rather than just a sequential number. This is extremely helpful when storing and cataloging the recordings. Some recorders have a computer audio-interface or the ability to connect a computer to record directly on a laptop or a tablet. This option allows the same recording quality while using special software for managing files (e.g., to tag files with a time-stamp and GPS position, or to automatically start and stop the recording according to received signals or according to a user-defined schedule).

### 2.3.5 Equipment for Monitoring Bats

Acoustic detection of ultrasonic bat calls has emerged as the most commonly used method for monitoring bat presence and activity (Collins and Jones 2009; Gorresen et al. 2008; Weller and Baldwin 2012). Observing and recording bats, other than for scientific research, is a very diffuse hobby and a common topic of citizen science. This results in a wide variety of bat detectors produced by small companies or DIY bat detector kits. The common types of detectors are heterodyne, frequency-division, time-expansion, zero-crossing, and full-bandwidth digital recorders (Obrist et al. 2010). Some bat detectors have their own specific software, either free or to be purchased, for further processing of recorded data.

Heterodyning was the first developed system, completely analog, to shift one frequency (the incoming signal) to another by multiplying it with a second frequency (set by the user). The user can tune the detector (similar to tuning a radio) to select a frequency range accessing a small portion of the available received frequency. For example, with a bat detector (e.g., Pettersson Elektronik[13] D100) tuned to the 40–50 kHz range, the call of a bat at 45 kHz (such as the Pipistrelli bat, *Pipistrellus* spp.) is multiplied (heterodyned) by a frequency (43 kHz) generated by an internal oscillator. This produces sidebands at 88 kHz and 2 kHz (which are the sum and the difference of the two frequencies); the higher frequency is eliminated with filters and the lower frequency is broadcast to the listener and available for recording. This makes for a tunable, inexpensive bat detector that will quickly indicate if bats are in the area. Heterodyning offers a limited view of the ultrasonic spectrum but is still appreciated by many bat specialists.

Frequency-division transforms the available frequencies and replicates the bat call by converting it into a square wave (sine wave also used) at its zero-crossing points. This wave is then divided by a preset factor (usually 10), creating another square (or sine) wave at a lower frequency (e.g., a 40-kHz call is converted to 4 kHz). All sounds in the environment are converted in this way. As such, masking of bat calls by noise, or overlapping of calls from different individuals, can produce results that could become difficult to interpret. Many devices have filters and ways to lower or otherwise adjust background noise. However, this recording option is now obsolete because modern digital ultrasound recorders are capable of recording at very high sampling frequencies (upward of 200 kHz) and capture the full bandwidth.

Time-expansion bat detectors use an AD-converter to digitize sounds, convert them so that they are audible to the human operator, and store these digital signals to memory (usually SD-card). Reduction of the recorded frequencies expands the sounds in time (hence the name). Some modern digital bat detectors do convert ultrasounds to audible sounds in real-time by means of FFT processing (Pavan et al. 2001). However, there is a delay when the signals are retrieved and played back at a slower speed (so that they can be heard with some delay). A high-frequency modulated call that sounds like a

[13] http://www.batsound.com/; accessed 15 Mar. 2021.

quick click is heard as a descending note or whistle upon playback from time-expansion.

Zero-crossing is an algorithm for extracting primary frequency information by tracking when the waveform crosses the zero-amplitude level at certain rates. Zero-crossing bat detectors run constantly, wake up when certain frequencies are detected, and save information on zero-crossings in storage. Some advanced bat detectors also retain the amplitude envelope of the original call; however, they only track the most intense component of the call. Using zero-crossing, a bat detector documents the dominant frequency, so if, for some reason, a harmonic is dominant over the fundamental or other signals overlap the fundamental of the call, only the most intense frequency is recorded. The operator needs to recognize this in order to represent the true nature of the bat's signal. The recordings produced by zero-crossing detectors are usually small (e.g., 50 KB), whereas an equivalent recording of full-spectrum calls consumes considerable storage space (e.g., 5 MB per call).

Full-spectrum digital bat detectors are digital recorders with high sampling frequency that capture the full bandwidth of the call (Dannhof and Bruns 1991; Moir et al. 2013). In some detectors, it is also possible to hear sounds in time-expansion while recording continuously. These bat detectors can record continuously or only when there are signals in a given frequency band set by the user (triggered recording); this solution reduces the storage size and shortens the time needed to analyze the recordings as only call series are recorded. Different trigger parameters allow selecting the frequency range to be recorded (spectral trigger) and the threshold level to activate the recorder. This technology is available in handheld and autonomous recorders (see Sect. 2.4.1), and computer-based bat detectors that use an external ultrasonic microphone. Some of the more advanced handheld digital bat detectors incorporate a display to visualize detected calls, and also include frequency-division, time-expansion, or frequency-shifting to provide acoustic feedback to the operator.

Some frequency-division detectors are combined with heterodyne and time-expansion capabilities into one unit. The Ciel CDB301 combines both a heterodyne detector with a frequency-division detector, allowing the researcher to tune into the frequency of a known bat call and identify a bat by both its sound contour and frequency. At the same time, the detector monitors the whole frequency band and checks if there are any bats in the vicinity. The Pettersson D240, like many of these dual bat detectors, provides heterodyning ability on one channel and time-expansion on another. Connected to a voice-activated digital recorder, these detectors can be left in the field in monitor mode and retrieved data can be analyzed on a PC using the product's software (e.g., *BatSound*). The Anabat Walkabout (Fig. 2.17) records bat signals using the zero-crossing technology and also saves signals as full-spectrum WAV files compatible with *SonoBat* software. The calls can be heard and displayed at the same time and saved to disk, making species identification instantaneous. Units are compact, mobile, and well-suited for long-term monitoring. Solar-powered units with detachable solid-state hard drives allow for greater periods of use.

For teaching or demonstration, any detector is useful, but one may consider heterodyne types of detectors because of their low cost (i.e., every student could use one). An interesting and flexible option is represented by ultrasonic microphones that incorporate a high-speed AD-converter that can be connected by USB to any computer platform (Windows, MacOS, Linux, iOS, Android, or Raspberry). The Dodotronic Ultramic series, the Wildlife Acoustics Echo Meter Touch, and the Petterson M500 are great devices for classroom demonstration. They allow to record ultrasounds continuously or on trigger with a companion tablet or smartphone, and provide full-spectrum recording capability, audio feedback, and real-time visualization. Some of these manufacturers also provide software for either basic operations, such as recording and display, or more advanced tasks such as bat species identification.




**Fig. 2.17** Some of the detectors discussed in this section. (**a**) Dodotronic USB Ultramic 384BLE, (**b**) Wildlife Acoustics (http://www.wildlifeacoustics.com/; accessed 15 Mar. 2021) Echo Meter Touch 2 Pro connected to an iPad and to a smartphone, (**c**) Anabat Walkabout (Titley Scientific (http://www.titley-scientific.com/; accessed 15 Mar. 2021)), and (**d**) D1000X bat detector by Pettersson Elektronik. Permission given by the respective manufacturers

### 2.3.6 Projectors

Playback studies to investigate animal behavior have been used on many different taxa (see Chap. 3, section on playback methods). The projectors used for broadcasting in air and under water also have, like the sensors, their characteristic frequency response and operational frequency range. Equipment with suitable characteristics should be chosen appropriately based on the characteristics of the sounds to be transmitted. Usually, speakers are electrodynamic devices; however, for high frequencies, electrostatic speakers are also used. At high amplitudes, projected sounds can distort. One must look in the manufacturer's manual to check maximum amplitude output of the projector and select a unit sufficiently capable of producing amplitude output similar to the level an animal would encounter. Generating sound in water requires more energy than in air, because of the higher impedance and density of water.

Among loudspeakers, some common names are used to describe their general operational frequency range: a tweeter is a high-frequency speaker typically small in diameter and a woofer is a low to very low frequency speaker that is much larger in diameter than a tweeter. A system with detachable loudspeakers can be convenient for placing speakers close to an animal or on opposing sides of an animal.

For underwater applications, there are two types of projectors: electrodynamic devices and transducers with piezoelectric elements. An electrodynamic device functions like an in-air speaker, but is watertight and can be used at



**Fig. 2.18** Photograph of JA Thomas lowering a Lubell underwater speaker into a melt hole to play back underwater vocalizations to Weddell seals (*Leptonychotes weddellii*) in the Antarctic

shallow depths. For example, a swimming pool speaker (Lubell,[14] Fig. 2.18) is an inexpensive electrodynamic device, but has a narrow frequency range that is relatively flat. On the other hand, piezoelectric projectors have projection sensitivity that varies with frequency. Note that many of the piezoelectric projectors are two-way or reciprocal devices that can also receive acoustic signals in water. The receiving sensitivity is fairly flat for a large portion of the operative frequency range; on the contrary, when working as a projector, the amplitude of the generated signal typically increases with frequency.

## 2.4 Autonomous Recorders

Autonomous recorders combine the different components of the signal chain (sound sensing, amplifying, filtering, and digitization) to offer a packaged solution. A variety of autonomous passive acoustic monitoring (PAM) systems have been developed, which allow the documentation of acoustic activity from animals and the environment. Autonomous recorders (both terrestrial and aquatic) are programmable and can be set up to satisfy specific needs. These systems can obtain long-term (months to years) data from remote areas and operate independent of weather and light conditions (e.g., Lammers et al. 2008; McCauley et al. 2017; Obrist et al. 2010). Some recorders generate recordings in popular formats (e.g., WAV files) that are compatible across several analysis software packages, whereas others generate a device-specific file format requiring the use of a specific software program for analyses. Autonomous recorders eliminate the influence of an observer's presence on the animal's behavior, are non-invasive, operate remotely, allow systematic periodic sampling, and provide long-term recordings.

### 2.4.1 Terrestrial Recorders

Autonomous recorders are used to study airborne sounds from terrestrial animals on a long-term basis, during day and night, during any type of weather, and in areas where the animals might not be visible because of vegetation. They are low-power, digital recorders with extended data storage capabilities enabling the recording of sounds for extended periods, continuously, or on a pre-defined schedule (e.g., record $x$ hours before and after sunset or sunrise, or for $x$ min every $y$ min). Important features of autonomous

[14] http://www.lubell.com/; accessed 15 Mar. 2021.

recorders in the field include: battery duration, total recording time, recorder reliability, programming capabilities, weatherproof construction, tamper-proof setup, ease of data-retrieval, and possible interface with video. The frequency response, dynamic range, and amplitude sensitivity of the unit are determined by the sound sensor, preamplifier, amplifier, and AD-converter used. By using a GPS or a highly precise internal clock, individual recorders can be time-synchronized. This allows measuring the TDOA of sounds among multiple recorders to triangulate and locate a sound source (see Chap. 4, section on localization). Another option is triggered recordings. For example, when the energy in certain frequency bands exceeds a preset threshold, data are recorded. This can reduce the amount of data to be stored onboard. Recorded data can be retrieved manually from the recorder or remotely via wireless methods. The more advanced units feature Wi-Fi, cellular network, or satellite communication interfaces for data transmission to a remote server. For instance, Pavan and team used autonomous recorders (Wildlife Acoustics SM3 and SM4) to document airborne sounds for six years at three locations with 10-min samples every 30 min (Fig. 2.19) (Pavan et al. 2015; Righini and Pavan 2019). Bat nocturnal activities were monitored via ultrasonic autonomous recorders (Wildlife Acoustics




**Fig. 2.19** (**a**) Photograph of autonomous acoustic recorders placed in the Sassofratino Nature Reserve, Italy. In the foreground, a Wildlife Acoustics Song Meter SM3. In the background, a custom recorder developed at the University of Pavia. (**b**) Wildlife Acoustics Song Meter SM4BAT-FS. (**c**) Titley Scientific Anabat Express. Permission to reprint by the respective manufacturers

EM3+ and SM4BAT-FS) and an ultrasonic USB microphone (Dodotronic Ultramic 250 K) connected to a PC-tablet.

The increasing interest in acoustic monitoring in the last few years has stimulated the development of many autonomous recorders; among these, the Wildlife Acoustics series, the Bioacoustic Audio Recorder (Frontier Labs,[15] Brisbane, Queensland, Australia), the Swift (Cornell Lab of Ornithology, Cornell University, Ithaca, New York, USA), and the Anabat Express (Titley Scientific, Brendale, Queensland, Australia). Some recent open-source examples are built around the Raspberry Pi and similar small-board computers. In some cases, the projects are open access. However, these devices often require large batteries to sustain power over long periods. Examples include the Solo acoustic monitoring platform[16] (Whytock and Christie 2017), based on the Raspberry Pi and an external microphone; the Bat Pi 2[17] for monitoring bats; and the AURITA system, which combines in a waterproof package the Solo recorder and a commercially available bat recorder, the Peersonic RPA2, to capture sounds from 60 Hz to 192 kHz (Beason et al. 2018). The AudioMoth,[18] an open-source device, which also can be purchased and assembled, employs a low-power microcontroller and an onboard MEMS microphone (Hill et al. 2018) and has very basic capabilities but allows remote data acquisition at very low cost on a single channel with sampling frequencies up to 384 kHz.

### 2.4.2 Underwater Recorders

Over the past few decades, interest in marine bioacoustics and in underwater noise monitoring have increased worldwide, and the market for underwater autonomous recorders is rapidly expanding. Autonomous recorders with a variety of features (such as operational longevity, high depth rating, onboard processing, and communication capabilities) are produced by several commercial organizations and academic entities. Examples of commercially available recorders are the AMAR from JASCO Applied Sciences,[19] Snap from Loggerhead Instruments,[20] AURAL from Multi-Électronique,[21] icListen from Ocean Sonics,[22] SoundTrap from OceanInstrumentsNZ,[23] EAR from Oceanwide Science Institute[24] (Lammers et al. 2008), and RESEA from RTSYS.[25] Academic recorders include the Rockhopper by Cornell Lab of Ornithology (upgraded variant of MARU; Klinck et al. 2020), USR by Curtin University (McCauley et al. 2017), and HARP by Scripps Institution of Oceanography (Wiggins and Hildebrand 2007). Selection of a particular type of autonomous recorder is driven by the needs and limitations of the research project. Most of these modern recorders support recording at 16- and 24-bit resolutions and offer flexibility to record at different sampling frequencies and to program custom duty cycles. Some even offer the flexibility to easily switch components (e.g., choosing hydrophones with appropriate sensitivity or frequency range). With the market for these recorders expanding, there are numerous options available beyond the few products mentioned here.

In very shallow waters, at depths reachable by a diver, deployment and recovery operations can be relatively easy. At greater depths, specific additional equipment is needed to allow the recovery—typically, a ballast (to secure stability on the seafloor), an acoustic release, and floaters to retrieve the recorder at the surface once the

[15] https://frontierlabs.com.au/; accessed 23 Aug. 2021.

[16] http://solo-system.github.io/home.html; accessed 15 Mar. 2021.

[17] http://www.bat-pi.eu/; accessed 23 Aug. 2021.

[18] https://www.openacousticdevices.info/; accessed 23 Aug. 2021.

[19] http://www.jasco.com/; accessed 15 Mar. 2021.

[20] http://www.loggerhead.com/; accessed 15 Mar. 2021.

[21] http://www.multi-electronique.com/; accessed 23 Aug. 2021.

[22] http://oceansonics.com/; accessed 15 Mar. 2021.

[23] http://www.oceaninstruments.co.nz/; accessed 15 Mar. 2021.

[24] https://oceanwidescience.org/; accessed 23 Aug. 2021.

[25] http://rtsys.eu/; accessed 15 Mar. 2021.

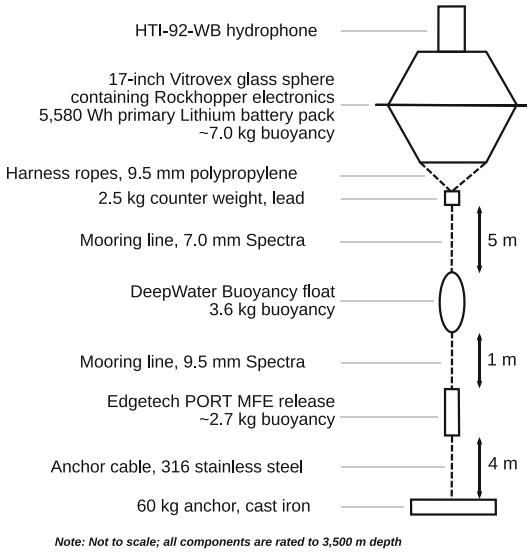



**Fig. 2.20** Schematic of a mooring setup for the Rockhopper autonomous passive acoustic recorder (Klinck et al. 2020). The example includes a wide-bandwidth hydrophone from HighTech Inc. (http://www.hightechincusa.com/; accessed 15 Mar. 2021) (HTI-92-WB), but the recorder offers flexibility with hydrophone choices

releaser disconnects the recorder from the ballast (Fig. 2.20). Anchored units are sometimes also diver-recovered or programmed to surface at a set date and time. In ice-covered habitats, the equipment can be secured to fast- or pack-ice with the hydrophone in the water.

## 2.5 Recording Directly to a Computer

Almost all computers, laptops, and tablets have an audio input and built-in microphone. Digital recording of sounds is controlled by the onboard soundcard. However, in most cases, the recording quality of the built-in microphone is only conducive for recording human voice or music and inadequate for animal sounds. For most animal recordings, an external sound sensor (microphone or hydrophone) connected to a high-quality audio input must be used with the computer or laptop. The recordist should consult the computer specifications to know the frequency range and dynamic range of the built-in soundcard. If the built-in sound system of a computer is not good enough, an external AD-converter can be easily connected by USB, or, for special devices, by other interface types. For fieldwork, it is preferable to choose converters with powering from the computer USB. The quality of recordings depends on the preamplifier noise and bandwidth, sampling rate, and bit-resolution of the soundcard or AD-converter. However, other features can drive the choice: number of channels, features of the AD-converter, the type of interface (USB, Firewire, Thunderbolt, or proprietary), availability of drivers for the computer, and power available for the sensors (P48 or PIP). For laptops used in fieldwork, their size, weight, ruggedness, power consumption, and reliability should be considered. Most USB-based converters for music recording are equipped with microphone preamplifiers with P48 power and offer good quality; some offer very high quality, comparable to the best digital recorder, with sampling frequencies up to 192 kHz with a number of channels ranging from 2 to 8; some external units provide up to 32 channels. Single-channel AD-converters are also available to be directly connected to a P48 microphone, to transform the microphone into a USB microphone. However, because some quality parameters are rarely described in official specifications (e.g., the self-noise, jitter-noise, and the anti-aliasing-filter used), conducting laboratory or bench tests to choose the best AD-converter can be necessary. For specific applications, the use of instrumentation AD-converters may be required.

## 2.6 Calibration

For quantitative animal bioacoustic studies, calibrated recording equipment needs to be used so that absolute sound pressure can be determined. This section deals with two types of calibration: calibrating the recording equipment and calibrating the recording. To calibrate the recording, the calibration of the recording equipment is applied to the recorded data.

Calibrating the recording system implies determining the frequency response and amplitude

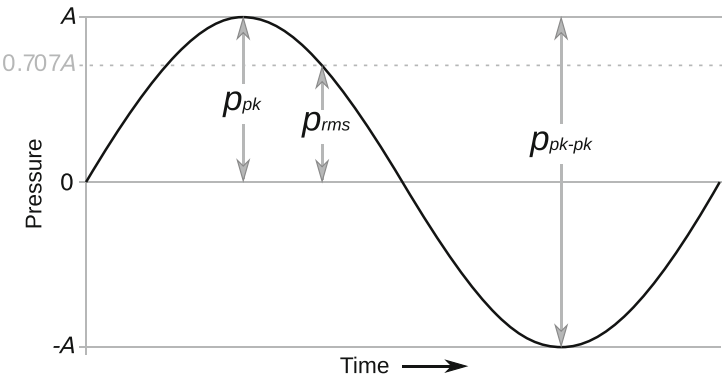



**Fig. 2.21** Waveform of a sinusoidal signal (pressure $p$ as a function of time) showing $p_{rms}$, $p_{pk}$, and $p_{pk\text{-}pk}$

sensitivity of the recording system. The recording system consists of several components (e.g., sensor, amplifier, and AD-converter), each with its own frequency response and amplitude sensitivity. The recording system may be calibrated as a whole by presenting a calibration signal of known amplitude and measuring the output. From the difference between output and input, the frequency response and amplitude sensitivity may be calculated. Or, each piece of equipment may be calibrated separately, and the frequency responses and amplitude sensitivities may be joined (i.e., multiplied in linear terms or summed in logarithmic terms).

The simplest calibration signal is a sine wave (i.e., a pure tone; Fig. 2.21). While the rms value is typically used in equipment calibration sheets, the peak (pk) or peak-to-peak (pk-pk) values are more easily read off signal displays on a computer or oscilloscope. For a sine wave, the conversion is:

$$p_{rms} = \frac{p_{pk}}{\sqrt{2}} \approx 0.707 \times p_{pk}$$

$$\Leftrightarrow 20 \log_{10} \frac{p_{rms}}{p_0} = 20 \log_{10} \frac{p_{pk}}{p_0} - 20 \log_{10} \left(\sqrt{2}\right)$$

$$\approx 20 \log_{10} \frac{p_{pk}}{p_0} - 3\,\text{dB}$$

The variable $p$ denotes pressure. The reference pressure $p_0$ is 20 μPa in air (i.e., for microphone calibration) and 1 μPa in water (i.e., for hydrophone calibration); also see Chap. 4 on an introduction to quantities and units. To add to the confusion, the dynamic range of analog electronics and AD-converters is given in pk-pk values. The simple equation is only valid for sinusoidal signals.

Using a sine wave yields an amplitude sensitivity at only one frequency. In order to measure the frequency response of the equipment, a series of sine waves at different frequencies needs to be presented. More commonly, white noise (i.e., a broadband signal of equal amplitude across frequency) is used and amplitude sensitivity is determined at all frequencies contained in the signal after Fourier transform of the output signal (see Chap. 4).

A simple recording setup is shown in Fig. 2.22. A calibration signal $p(t)$ (i.e., pure tone or white noise of known amplitude) is presented to the sensor (i.e., microphone or hydrophone). The sensor has a sensitivity $s$, which relates the voltage $V$ at its output to the pressure $p$ at its input; so $s$ has the unit V/Pa. The sensitivity can also be expressed in dB re 1 V/Pa: $S = 20 \log_{10} (s/(\text{V/Pa}))$. The output voltage $V$ of the sensor is typically passed to an amplifier. The amplifier gain $g$ relates the voltage at its output to the voltage at its input and is thus unit-less: $g = V_2/V_1$. Expressed in dB, the amplifier gain is $G = 20 \log_{10} (g)$. The output voltage of the amplifier is then passed to an AD-converter such as a soundcard on a computer. The AD-converter has a digitization gain $c$, that relates the digital values $d$ in the audio file to the voltage $V$ at its input. The bit-depth of the AD-converter limits the maximum digital value (i.e., the full-scale value $FS$) that can be stored. The digitization gain is defined as the ratio of the full-scale value

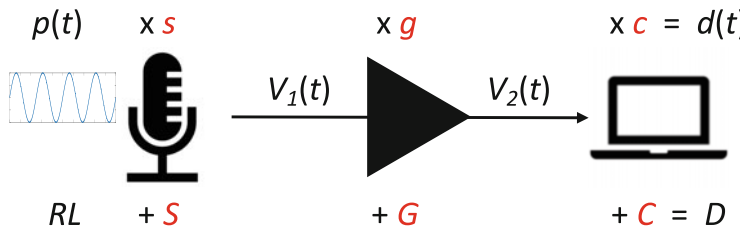



**Fig. 2.22** Sketch of a generic recording system consisting of a sensor (i.e., microphone or hydrophone), amplifier, and AD-converter (e.g., a computer with soundcard). Each piece of equipment has its own sensitivity or gain (indicated by red letters). These sensitivities may be expressed in linear terms (small letters) or decibels (capital letters). The sensor converts the input pressure time series $p(t)$ to a voltage time series $V_1(t)$, which is amplified to yield $V_2(t)$. The AD-converter produces a digital time series $d(t)$

to the input voltage that produces the full-scale value: $c = FS/V_{max}$. The digitization gain is expressed in dB re FS/V. The sensitivities (in linear terms) of each component in the recording system can be multiplied to yield the system sensitivity, which relates the digital values $d$ in the audio file to the pressure $p$ sensed by the sensor. In logarithmic terms, the overall system sensitivity is the sum of the sensitivities of each piece of equipment.

Once the recording system has been calibrated, it can be used to record animals or other sound sources. To determine the calibrated pressure time series $p(t)$ from the stored data $d(t)$, divide by all the sensitivities and gains: $p(t) = d(t) / (c\ g\ s)$. Alternatively, using the level quantities (in dB) for each equipment, the received level $RL$ (e.g., rms sound pressure level) is determined by subtracting all sensitivities and gains from the rms amplitude level $D$: $RL = D - C - G - S$. For example, somebody made a 10-minute recording of a singing bird. The microphone sensitivity was $s = 50$ mV/Pa, or $S = 20\log_{10}(0.05) = -26$ dB re 1 V/Pa. The amplitude at the output of the microphone was amplified by, let's say, a factor $g = 100$, or $G = 20\log_{10}(100) = 40$ dB. The soundcard produced a full-scale amplitude at 2 V input: $c = FS/2$ V, or $C = 20\log_{10}(1/2) = -6$ dB re FS/V. A computer is used to process the data. If the data are read using the *MATLAB* (The MathWorks Inc., Natick, MA, USA) function *audioread* with the flag "*native*," then the raw digital values are presented. With the flag "*double*," the data are normalized by the full-scale value and so lie between −1 and +1. Computing the rms amplitude of the normalized digital time series yields a value of, let's say, 0.06. In logarithmic terms, the rms amplitude level of the stored normalized data is $D = 20\log_{10}(0.06) = -24$ dB. What was the received sound pressure level of the bird song? Subtracting all the gains, the rms sound pressure level received at the microphone was −32 dB re 1 Pa (because −24 –(−6) – 40 –(−26) = −32). The standard reference pressure in air is, however, 20 μPa, which is equivalent to $20\log_{10}(20/1{,}000{,}000) = -94$ dB re 1 Pa. So, the rms sound pressure level recorded from the bird was −32 −(−94) = 62 dB re 20 μPa. The researcher might further want to compute calibrated sound spectrograms of the bird song, and so the question is how to convert the digital values to pressure values. Using the linear sensitivities and gains, $p(t) = d(t) / (FS / 2\ \text{V}) / 100 / (0.05\ \text{V/Pa})$ yields pressure samples in units of Pa.

### 2.6.1 Microphone

To make accurate recordings of sound intensity in the laboratory or field, either from an animal or a different source, a researcher should always use a calibrated microphone. A commercial microphone is calibrated when received from the manufacturer and comes with specification sheets containing amplitude sensitivity, frequency response, and reception directionality as a

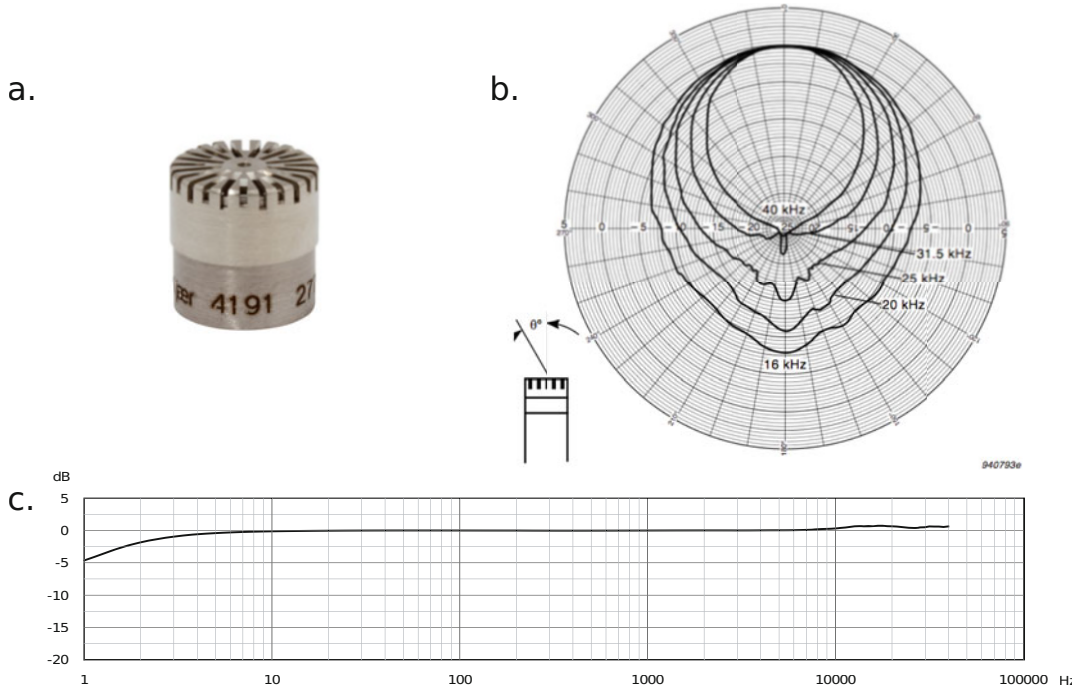



**Fig. 2.23** Specifications of a Brüel & Kjær 1/2-inch free-field microphone type 4191. (**a**) Photo. (**b**) Polar plot of receiving directionality from 16 kHz to 40 kHz. c. Graph of frequency response. Permission to reprint from Brüel & Kjær

function of frequency in the horizontal and vertical planes. For example, the ½-inch microphone shown in Fig. 2.23a has an amplitude sensitivity of 12.5 mV/Pa or −38 dB re 1 V/Pa and a flat frequency response (to within 3 dB) from about 3 Hz to 40 kHz (Fig. 2.23c). Given its cylindrical symmetry, it is omnidirectional about its vertical axis (Fig. 2.23b). In the vertical plane, its receiving directionality is steered toward its axis; in other words, it is most sensitive in the forward (i.e., vertical in Fig. 2.23b) direction. The lower the frequency, the more receptive it becomes from other directions. To check that the microphone maintains its sensitivity over time, a bioacoustician should periodically use a calibrator. For example, the calibrator shown in Fig. 2.24 is very stable and emits a 1 kHz tone at 94 dB re 20 μPa.

Provided there is a commercial, calibrated microphone available, a researcher can calibrate a microphone of unknown sensitivity by comparison with a calibrated microphone. Using a loudspeaker system to do this is a convenient option. Alternatively, signals of opportunity, like roadway or jet noise, may also be considered while ensuring that both microphones receive the same signals and levels. First, calibrate the sound field at the frequencies of interest with the calibrated microphone. Then, replace the calibrated microphone with the one of unknown

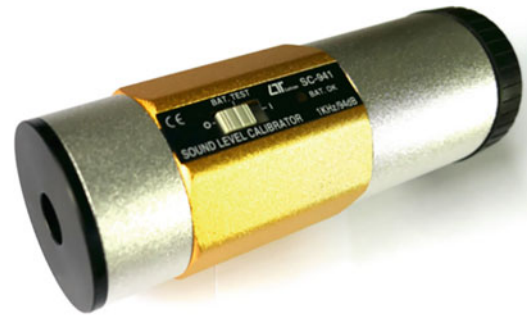



**Fig. 2.24** A sound level calibrator (LUTRON, model SC-941) that generates 94 dB re 20 μPa at 1 kHz. The microphone to be calibrated must be inserted in the hole (1/4 inch diameter) on the left side. Adapters are available to fit other microphone diameters

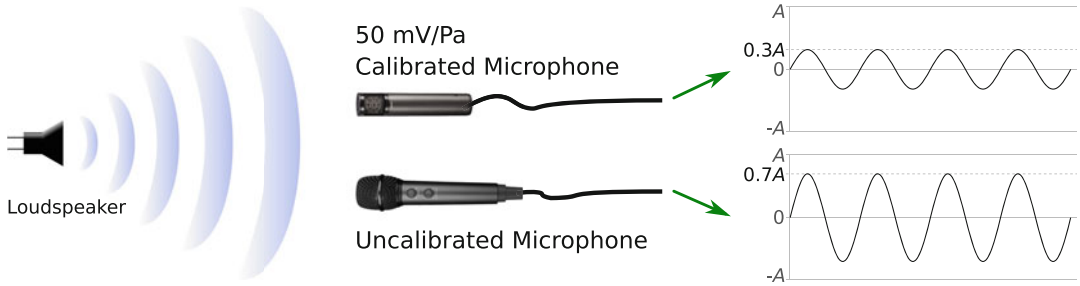



**Fig. 2.25** Sketch of a setup to calibrate a microphone of unknown sensitivity with a microphone of known sensitivity in a constant sound field. Redrawn from a laboratory manual with permission from Lasse Jakobsen, Institute of Biology, University of Southern Denmark, Odense, Denmark

sensitivity and record the output in the same frequency range. Do not place the two microphones side-by-side in the sound field since this could cause diffraction and distortion of the sound field. The sound field should not contain echoes, so choose an open space or an anechoic room for low frequencies. In the example of Fig. 2.25, the calibrated microphone has a sensitivity of 50 mV/Pa. In the given sound field, it produces an output signal with an amplitude of 0.3 voltage units. After the calibrated microphone has been removed and the to-be-calibrated microphone has been installed at exactly the same location, the latter produces an output signal of 0.7 voltage units. The sensitivity of the to-be-calibrated microphone is simply $0.7/0.3 \times 50$ mV/Pa $= 117$ mV/Pa.

### 2.6.2 Hydrophone

High-quality commercial hydrophones are calibrated by the manufacturer with all pertinent information contained in the accompanying specification sheets. Many hydrophone types have built-in preamplifiers with amplification and impedance matching. Thus, these hydrophones come with a calibration sheet having one sensitivity value that includes the preamplifier. The sensitivity of a hydrophone is usually expressed in dB re 1 V/μPa, which is different from the expression for microphone sensitivity (dB re 1 V/Pa).

To use RESON hydrophones as examples, their most sensitive hydrophone (i.e., the one with the least negative sensitivity: TC4032; Fig. 2.26) has a sensitivity of −170 dB re 1 V/μPa (single ended). If the sound received by the hydrophone were 170 dB re 1μPa rms, then the output from the hydrophone would be 1 V rms. To compare this to a microphone, add 120 dB, which is a factor $10^6$ in pressure ($20 \log_{10}(10^6) = 120$ and $10^6$ μPa = 1 Pa). So, −170 dB + 120 dB yields −50 dB re 1 V/Pa. The most sensitive ½- or 1-inch microphone is −26 dB re 1 V/Pa, which is 24 dB (i.e., about 16 times, because $20\log_{10}(16) = 24$) more sensitive than the TC4032 hydrophone.

Although most hydrophones are stable through time, it is wise to check the calibration periodically using a pistonphone. However, a pistonphone can determine the sensitivity of an uncalibrated hydrophone at only one frequency. The sound pressure of a pistonphone is extremely stable and is only affected by one factor: barometric pressure. For this reason, a special barometer is included with the pistonphone. For accurate calibrations, the barometric pressure should be checked, and sound pressure adjusted according to the scale on the barometer. For calibrations performed near sea level (as is often the case in marine bioacoustics), this error is negligible, but if one is working in an aquatic environment that is significantly above sea level, then this factor (which is −2 dB at 2000 m altitude) should be included. For hydrophones to be deployed at

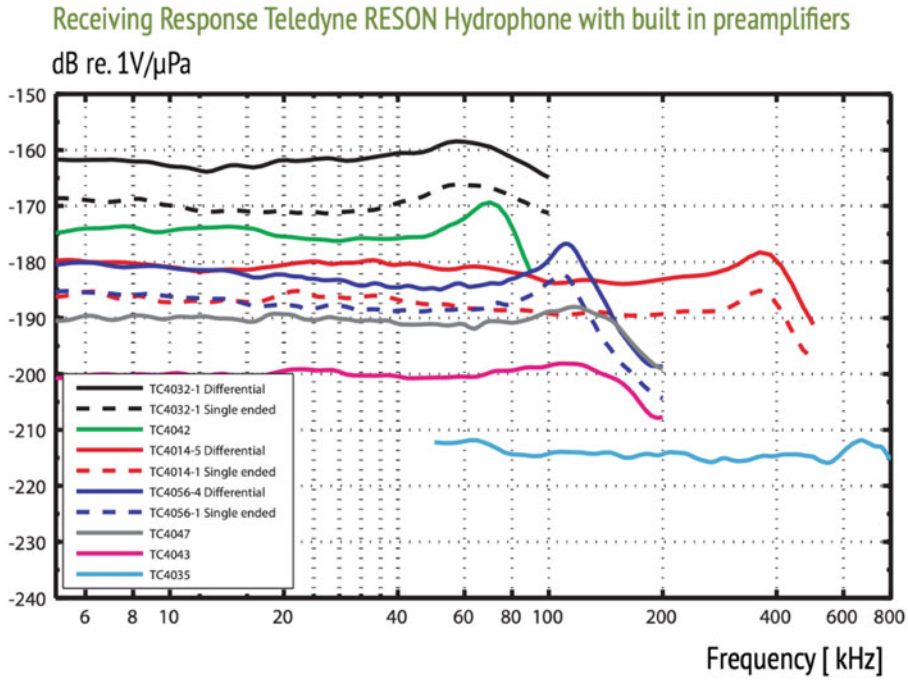



**Fig. 2.26** Graph of amplitude sensitivity and frequency response for several RESON hydrophones with preamplifiers. The most sensitive is the TC4032; the least sensitive is the TC4035. Permission to reprint from RESON (http://www.teledyne-reson.com/; accessed 15 Mar. 2021)

great depth in the ocean, the amplitude sensitivity (and pressure resistance) should be measured in a pressure chamber.

The frequency response of an uncalibrated hydrophone (for frequencies up to a few kHz) can be measured in air by using the same method as described for a microphone (Fig. 2.25). However, for higher frequencies, this should be done in open water (e.g., a deep lake) and the method described for microphones can be used by simply substituting the microphone with a hydrophone of known sensitivity compared to one of unknown sensitivity. An appropriate amplifier and an underwater projector are needed, but a hydrophone without a built-in preamplifier also can be used as a projector. First, the environment (lake, pool, or tank) should be checked for echoes and reverberations (see Popper and Hawkins 2018 for details). The projected calibration sound must be a pulse that ends before the first echo arrives at the sensor. This necessity restricts the frequency range that can be used for calibration since the projected pulse must be ramped up and down to reduce high-frequency artifacts caused by the onset and end of the pulse.

The next step is to determine the received level of an underwater sound. For example, a dolphin click is recorded with a TC4035 hydrophone, which has a sensitivity of −215 dB re 1 V/µPa (Fig. 2.26). If the output is amplified by 60 dB (1000x) and the recorded signal is 1.2 V pk-pk, then the received level is: 20 $\log_{10}$ (1.2) – 60 – (−215) = 1.58 − 60 + 215 ≈ 157 dB re 1 µPa pk-pk. Usually, the analog voltage signal is converted to a digital signal by an AD-converter, which has a digitization gain that also needs to be accounted for (see above).

### 2.6.3 AD-Converter

A 16-bit AD-converter has $2^{16}$ bit resolution, covering 65,536 counts peak-to-peak. Its full-scale value is $2^{16}$–1 = 65,535 in unipolar mode,

where the digital amplitude values lie between 0 and 65,535, or $2^{15} = 32{,}768$ in bipolar mode, where the digital amplitude values are in the range −32,768; . . ; 0; . . ; 32,767. In decibels, the dynamic range of a 16-bit AD-converter in bipolar mode is $20 \log_{10}(32{,}768) = 90$ dB. Every bit gives ~6 dB of dynamic range in the digital domain. But a 90-dB dynamic range rarely can be realized since most electronics used before AD-conversion do not have such a large dynamic range. A 24-bit converter in bipolar mode offers a theoretical dynamic range of about 138 dB; however, only the most sophisticated electronics can provide up to 115–120 dB of dynamic range. This means that there cannot be more than 19–20 bits of real dynamic range and the remaining bits (least significant bits) are just filled by noise. AD-converter specification sheets rarely show this, thus there is growing need to have more realistic AD-specifications to account for the intrinsic AD-converter noise and its artifacts showing as distortion and jitter. In some recording systems, the least significant bits are used to encode complementary information; however, this practice is not standard.

AD-converters thus carry an intrinsic digitization gain, which is the ratio of the full-scale value to the input voltage that leads to full-scale. The digitization gain is expressed in dB re FS/V. For example, an AD-converter with a digitization gain of −6 dB re FS/V reaches its FS value at a peak input voltage of 2 V, because $20 \log_{10}(\text{FS}/2\text{ V}) = -6$ dB re FS/V. AD-converters may be calibrated with a voltage signal generator. The peak voltage of the input signal has to be less than the maximum voltage range specified in the specification sheet; otherwise, the AD-converter will be overloaded and the signal clipped.

### 2.6.4 Autonomous Recorder

Off-the-shelf autonomous recorders are manufacturer-calibrated. The specification sheets typically give one overall amplitude sensitivity and frequency response for the entire system (including sensor, amplifier, and AD-converter). If the recorder allows variable gain settings, then the chosen gain will affect the amplitude sensitivity and needs to be accounted for. Some manuals (e.g., the SoundTrap User Guide[26]) provide guidance on how to calibrate the recorded data if read by software packages such as *MATLAB*, *PAMGuard*, or *Audacity*.

### 2.6.5 Measuring Self-Noise

When intending to record quiet sounds or ambient sound levels in the absence of nearby sound sources, it is important to first measure the system self-noise to avoid confounding electronic noise with environmental noise. For this, the system should record in a quiet room and the sound sensor should be in a sound- and vibration-proof box (Fig. 2.27). If using an autonomous recorder, the entire system should rest in a sound-proof box.

To record quiet sounds under water or to accurately quantify ambient sea noise, a sensitive hydrophone with a wide frequency range is needed (e.g., the TC4032, Fig. 2.26). All of the system components should have low self-noise. A "wet-ground" ground-wire from the input equipment to the water might be necessary to reduce system noise. The amplifier should have an adjustable band-pass filter to avoid aliasing during direct digital recording. The AD-converter needs sufficient bit-resolution and sampling rate to cover the frequency band of interest. The system frequency response shown in Fig. 2.27 goes up to about 100 kHz. If the full bandwidth is desired, then the sampling frequency should be at least 200 kHz. When reporting measured levels, provide the frequency range over which sound was measured and the bandwidth over which sound levels were computed (e.g., per Hz or in 1/3-octave bands).

[26] http://www.oceaninstruments.co.nz/wp-content/uploads/2018/03/ST500-User-Guide.pdf; accessed 5 Mar. 2021.

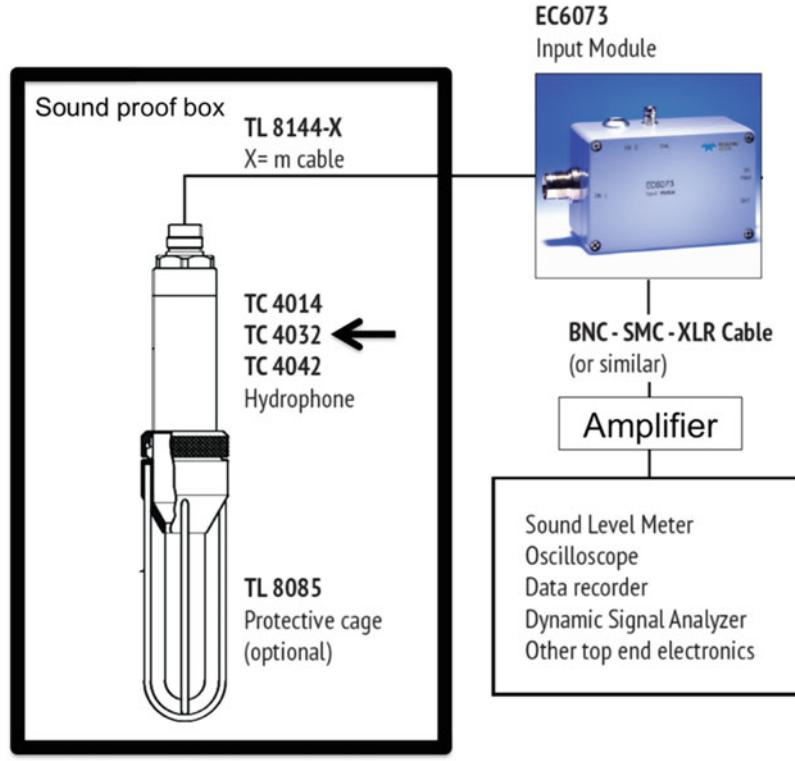



**Fig. 2.27** Diagram of equipment to measure underwater ambient noise. The RESON hydrophone with lowest self-noise is the TC4032. Prior to deployment, system self-noise may be determined by recording with the hydrophone in a sound- and vibration-proof box in the laboratory. Permission to reprint from RESON

## 2.7 Other Gear

### 2.7.1 Sound Pressure Level Meter

SPL meters, also called phonometers, are used to measure ambient noise, including abiotic and biotic sounds. SPL meters have a variety of settings for transient vs. continuous sound, frequency range, amplitude range, and any weightings (Brüel and Kjær 2001). The microphone on an SPL meter is omnidirectional, can be covered with a windsock, and mounted on a tripod. The fast-setting is used for impulse or transient sounds. The slow-setting is used for continuous sounds. Most SPL meters have a selectable frequency range. The user can select a flat setting, which collects dB measurements equally over the desired bandwidth (i.e., without weightings). The A-weighting is selected when the user desires to place a filter over the sampled frequency range in an effort to account for the relative loudness perceived by the human ear (see Chap. 4, section on weighting curves). However, it is important to not underestimate the impact of infrasounds, which can be heard or perceived by animals. The C-weighting is selected when the user desires to measure the peak sound pressure level. Measurements with these filters are expressed as dB(lin), dB(A), or dB(C). To measure environmental noise over the whole spectrum (especially for species with unknown hearing curves), it is important to use the unweighted, flat setting. At low frequencies of anthropogenic noise, the type of weighting used can make a large difference in the amplitude measurement.

Out of the various measures an SPL meter may report, the most common one is perhaps the Equivalent Continuous Sound Level ($L_{eq}$), which is a time-average: the equivalent constant SPL that would produce the same energy as the fluctuating sound level measured over a given time interval (e.g., 60 s). The duration of the measure must be declared as $L_{eq,T}$ (e.g., $L_{eq,60s}$),

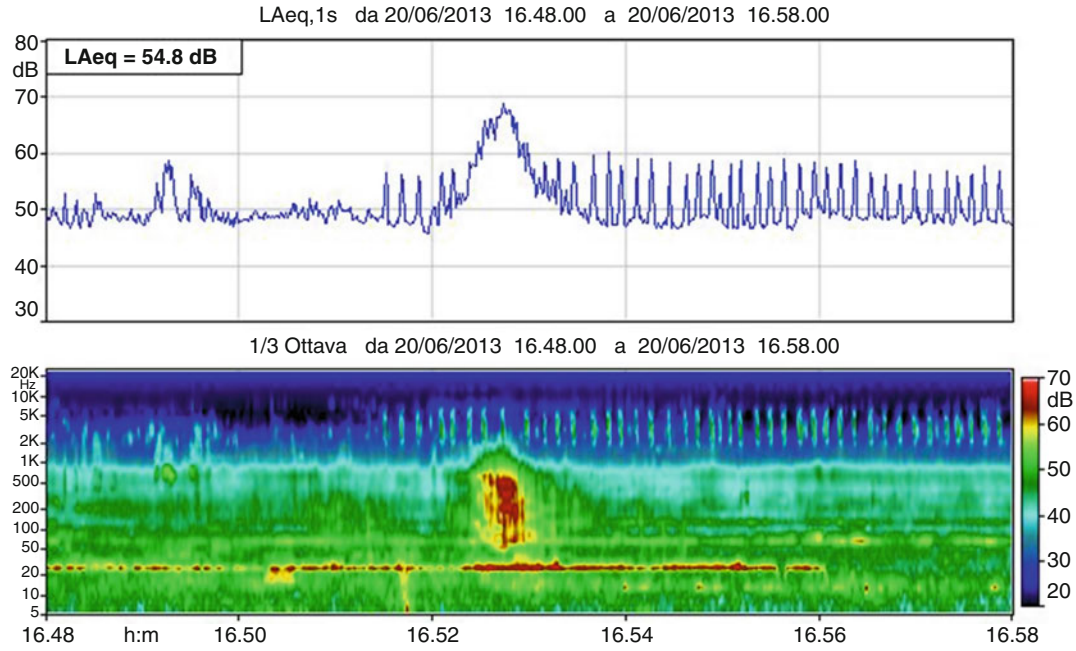



**Fig. 2.28** Recording and spectral analysis of noise in a residential area. Recording (top) of the overall sound level (A-weighted) with the $LA_{eq}$ level of the shown period. The unweighted spectrographic image (bottom), with frequency up to 20 kHz on a logarithmic scale, shows the spectral composition of the recorded period. At about 20 Hz is the noise generated by a truck engine. At about 16.53 occurs the noise of a passing airplane (50–1000 Hz). Bird songs appear at 1500–9000 Hz. Courtesy of Alberto Armani

where T is the time interval of the measurement. The level may be weighted (e.g., A or C weighting). $LA_{eq}$ is often used in the assessment of noise dose or sound exposure in humans (Fig. 2.28). For example, $LA_{eq,1s} = 73$ dB or $L_{eq,1s} = 73$ dB(A) is a measurement taken with an A-weighting filter over 1 s and $LC_{eq,1s}$ indicates a measurement taken with a C-weighting filter for 1 s.

Some SPL meters have a 60-s $L_{eq}$ setting used for short-term sampling. However, if the sound level varies randomly, calculating $L_{eq}$ is tricky, and so, Integrating Sound Level Meters are better (Fig. 2.29) as they determine $L_{eq}$ during a suitable time period. When more information on the statistics of sound levels is needed, in both time and frequency, noise-level analyzers are used (Fig. 2.29). They perform statistical analyses of sound levels over a specified period, either broadband or band-limited (e.g., in a 1-octave or 1/3-octave band). Most sophisticated, and expensive, noise measuring systems can produce spectra in narrower bands (as fine as 1-Hz bands) and calculate spectral percentiles to show the level variation statistics for each frequency band. In other words, the percentile analysis of a 1/3-octave spectrum shows what percentage of time each level is reached or exceeded within the measurement period (see Chap. 4, section on power spectral density percentiles).

All these devices need to be calibrated periodically with a known calibration tone. Calibrators are standardized at the factory and usually maintain calibration for a long time. Only specialized laboratories can certify calibrators. The calibrator signal is usually a 1-kHz sinusoidal tone at 94 dB re 20 μPa SPL rms (equivalent to a pressure of 1 Pa rms, 95.45 dB pk, or 1.41 Pa pk).




**Fig. 2.29** Photograph of Larson Davis SoundAdvisor 831C sound level meter with spectral analysis and sound recording capabilities (left; permission to reprint from Larson Davis (http://www.larsondavis.com/; accessed 5 Mar. 2021)) and of a simple noise-level analyzer with calibrator (right; shown being calibrated using a 1 kHz tone with 94 dB SPL)

## 2.7.2 Vibration Measurement

### 2.7.2.1 In Terrestrial Studies

In addition to communicating through sound (i.e., pressure waves propagating through air or liquid), animals ranging from elephants to insects communicate by producing waves that travel through solids (i.e., substrate-borne vibrations, also referred to as vibrational or seismic communication in the literature) (Cocroft et al. 2014a; Hill 2008; Hill et al. 2019; O'Connell-Rodwell 2010). Of insects alone, an estimated ~195,000 species communicate in part or whole via substrate-borne vibrations (Cocroft and Rodríguez 2005). Of these, the most species-rich group is plant-living insects, and so most examples in this section deal with invertebrate signalers and plant substrates.

Vibrational signals travel through various kinds of substrates (e.g., rod-like, such as plant stems; plate-like, such as leaf litter) as different types of waves (e.g., bending, Rayleigh) that vary in their direction of energy propagation (reviewed in Elias and Mason 2014; Mortimer 2017). In plant stems and leaves, substrate-borne vibrations travel as bending waves (Michelsen et al. 1982) and signal propagation is frequency-dispersive; in other words, energy at higher frequencies propagates faster than does energy at lower frequencies (Michelsen et al. 1982). Furthermore, each substrate acts as a unique filter, attenuating some frequencies more than others (reviewed in Elias and Mason 2014). Filtering varies among different plant species (Bell 1980; McNett and Cocroft 2008; Virant-Doberlet and Čokl 2004), different parts of same plants (Čokl et al. 2005; McNett and Cocroft 2008), and even among different parts of the same leaves (Čokl et al. 2004; Magal et al. 2000).

Filtering is a key consideration for selecting a sensor for recording or playback (Cocroft et al. 2014b). Importantly, the transmission and filtering properties of a given substrate can be affected by a sensor, if it loads on extra mass. If the aim is to characterize signal parameters of a given species, then to minimize filtering, one must choose a sensor that adds as little mass as possible and minimize the signal propagation distance between the source and the receiver. For example, one might affix a small and lightweight micro-accelerometer to the substrate, close to the signaling animal. Alternatively, one might use a laser-Doppler vibrometer to detect and record signals directly from the body of the signaling animal (Čokl et al. 2005).

The output of a sensor is proportional to the quantity (displacement, velocity or acceleration) that it detects – a sensor that detects displacement will be most sensitive to low-frequency signals, whereas a sensor that detects acceleration will be most sensitive to high-frequency signals. The consequence of this relationship between output and quantity is that the type of sensor used

impacts the measurements that one makes of a signal and how that signal is characterized.

Some of the key considerations for selecting a type of sensor include its sensitivity and power needs (all sensors require power), the frequency and amplitude ranges of the signals, equipment ruggedness and portability (if considered for fieldwork), and cost (Table 2.1). Research questions can be framed around the signaler or receiver, and the measurement of interest can vary widely (e.g., number of signals produced, signal parameters, etc.). Different sensor types function best in different frequency ranges, and the dominant frequency of a vibrational signal can vary widely, from <50 Hz for tremulating katydids (De Souza et al. 2011; Morris 1980; Morris et al. 1994; Sarria-S et al. 2016), to between 50 and 200 Hz for tremulating stinkbugs (reviewed in Čokl et al. 2014), to above 500 Hz for diverse kinds of plant-feeding insects (reviewed in Čokl et al. 2014). Vibrational signals can also be narrowband (McNett and Cocroft 2008) or broadband, with energy distributed over several kHz (Cocroft 1996; Hamel and Cocroft 2019).

The amplitudes of vibrational signals also vary widely, even just within small arthropods. For example, large neotropical katydids produce substrate-borne vibrations by vertically oscillating their abdomens relative to the substrate (in other words, they bounce) and the amplitude of these oscillations can be large enough to observe with the naked eye (Belwood and Morris 1987; Morris et al. 1994; Rajaraman et al. 2015). In contrast, the amplitude of signals by tiny treehopper nymphs can be so low as to be difficult to detect without a very sensitive sensor, such as a laser-Doppler vibrometer (LDV) (JH, pers. obs.). The animal's use of substrates is another key factor to consider: some vibrationally signaling animals, such as small, plant-feeding insects, are relatively sessile and signal from specific locations on plants of a single species (McNett and Cocroft 2008), whereas other vibrationally signaling animals are more motile and may signal on diverse substrate types (reviewed in Elias and Mason 2010).

### Sensor Types Based on the Quantity Measured

*Displacement*: Phonocartridges and other piezoelectric sensors have greatest sensitivity at low frequencies. Phonocartridges can be quite good for detecting low-frequency, low-amplitude signals in plant substrates, but placement of the photocartridge on the plant leaf or stem necessarily loads the substrate and changes its transmission properties (Fig. 2.30a). Additionally, amplitude measurements made with phonocartridges are variable and not repeatable, because amplitude varies with the pressure with which the stylus contacts the plant tissue.

*Velocity*: LDVs use the reflection of a laser beam pointed at a reflective object or substrate to detect the velocity of its movement. (If a surface does not reflect enough of the laser for measurement, a small amount of reflective paint or tape can be applied to the substrate.) LDVs are highly sensitive and excellent for detecting and making measurements of low-amplitude signals that also have energy concentrated in low frequencies. They do not load any mass to a substrate, so they do not affect signal transmission in this way, and in fact, they can be used to characterize signals by recording from an animal itself (Čokl et al. 2005). LDVs provide repeatable measures of amplitude for vibrational signals. Unfortunately, LDVs can be expensive. Although they are fairly portable, they are still quite cumbersome compared with a micro-accelerometer. Additionally, because an LDV detects motion perpendicular to the laser, the researcher must decide which plane is of interest (e.g., identify the major axis of motion). LDVs are not well-suited for high-amplitude signals, as a moving branch or stem will break the contact of the laser with the reflective surface and disrupt measurement.

*Acceleration*: Accelerometers can be purchased in a wide variety of sensitivities, frequency ranges, and sizes, and some models have the capacity for adjustable gain. For example, a commonly used micro-accelerometer in studies of small insects has a mass of 0.8 g and a frequency range of 0.8 Hz–10 kHz. Accelerometers can generate repeatable measurements of amplitude,

**Table 2.1** Examples of sensors that might be selected for vibrational communication studies, taking research aim, substrate type, signal frequency and amplitude ranges, ruggedness of equipment, and cost constraints into account

| Organism | Aim | Substrate type | Dominant frequency and range | Signal amplitude | Other considerations | Cost constraints | Decision |
|---|---|---|---|---|---|---|---|
| Small insect that mainly signals on herbaceous plant species (e.g., stinkbug) | Characterize spectral and temporal features of signals | Plant stem/leaf | 100 Hz, most energy <200 Hz | Low | Study will be conducted in the lab and field; field work is in moderate environments | Funding is provided, or an LDV is otherwise available | LDV is ideal for low-frequency, low-amplitude signals; will not load the substrate or affect signal transmission |
| Small, sessile insect that signals on a specific part of a single plant species (e.g., treehopper) | Behavioral study documenting signaling response (or lack thereof) in response to a stimulus | Tree branch | Broadband, energy distributed 100 Hz −5000 Hz | Low | Equipment needs to be sturdy; field study; branches sometimes move in breeze | Funding is limited | Accelerometer is the best bet for funding constraints, field study, and the need to affix sensor to substrate; frequency range suggests that this approach will work |
| Medium or large, motile arthropod (e.g., wolf spider) | Characterize spectral and temporal features of signals | Leaf litter | Both high and low-frequency signals | Medium; multiple signal elements | Animal moves around but is confined to an arena; signals on leaf litter | Funding is provided, or an LDV is otherwise available | LDV is ideal for not loading the substrate and for characterizing low- and mid-frequency ranges |
| Large, motile insect (e.g., katydid) | Behavioral study documenting signaling response (or lack thereof) in response to a stimulus | Large plant stem/leaf | 20 Hz, most energy < 100 Hz | High (can observe movement of plant stem as animal signals) | Equipment needs to be sturdy and function in high heat and humidity environment | Funding is limited | LDV is ideal for low-frequency signals, but high signal amplitude suggests that laser would not remain in contact (with a moving substrate). Funding limitations and study setting suggest that an accelerometer is the place to start. |

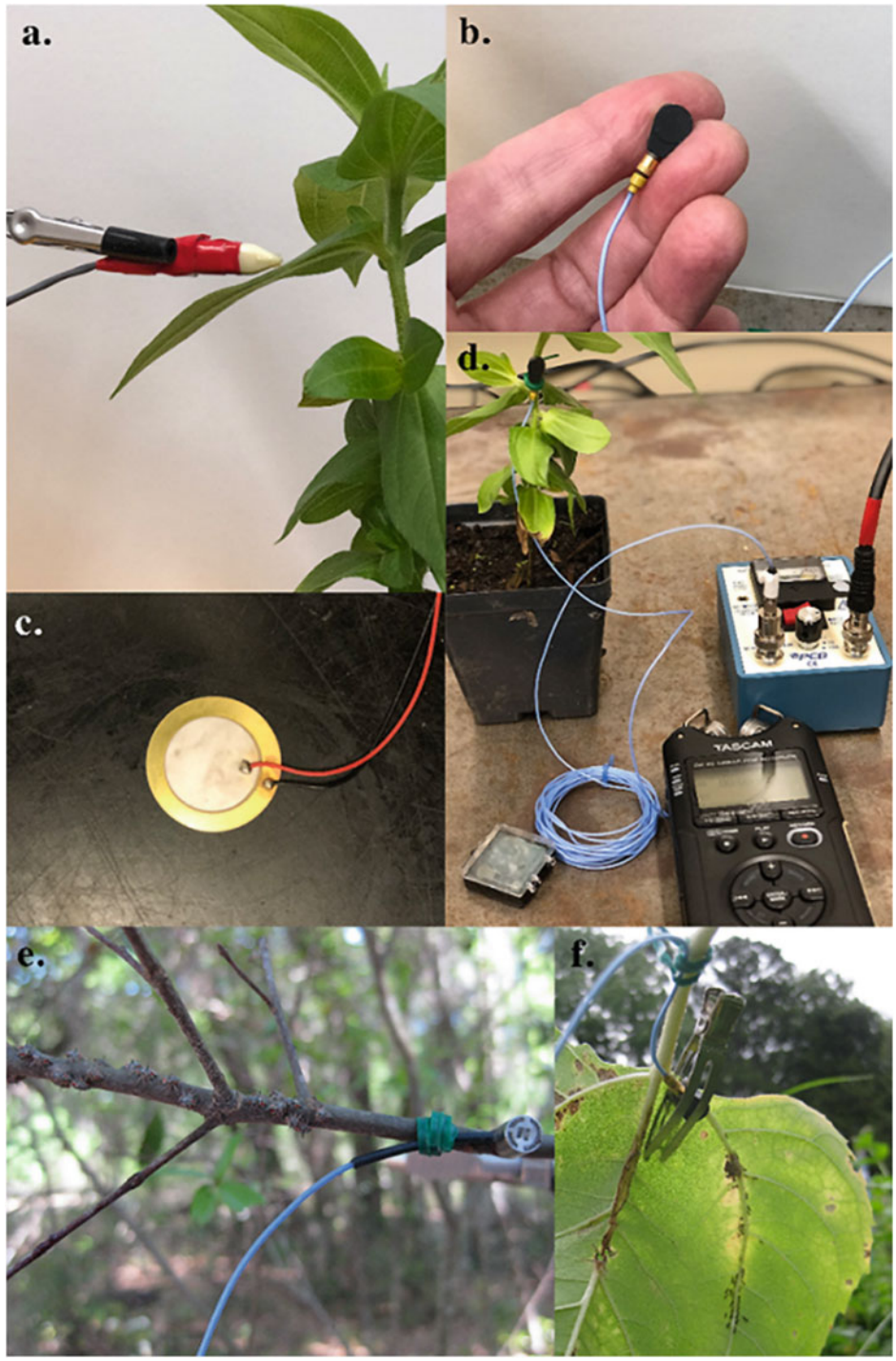



**Fig. 2.30** Sensors that detect and measure substrate-borne vibrations. (**a**) A phonocartridge attached to lab-hands or a thin wooden dowel. (**b**) Accelerometer. (**c**) Piezo disc or contact microphone for detecting substrate-borne vibrations. (**d–f**) Accelerometers affixed to substrates with a small amount of accelerometer wax or dental wax. Lightweight supports such as twist-ties and thin hair clips are used to reduce the likelihood of the accelerometer shifting position or detaching from a substrate

and because accelerometers are necessarily attached to a substrate, they can measure high-amplitude signals that move the substrate itself. Accelerometers are lightweight and small (Fig. 2.30b), can be rugged, and several commonly used models can be powered by one or more 9-V batteries. Drawbacks of accelerometers are that attaching a sensor to a substrate loads mass to the substrate; to avoid altering of substrate transmission properties, it is recommended to limit sensor mass to <5% of the mass of the substrate (Cocroft and Rodríguez 2005). Because accelerometers detect acceleration, they are not as sensitive at low frequencies as they are at higher frequencies, and they generally have lower bandwidths than LDVs.

The study of animal vibrational communication is rapidly growing. In order to withstand the rigor of peer-review, researchers must document the type, make, model, and sensitivity of the sensors used, and also document the factors likely to affect signal characteristics and propagation (e.g., substrate type and characteristics, position of the animal). The relative position of the sensor must be logical, consistent, and be informative for the study. For sensors that attach to substrates (e.g., accelerometers), secure and even attachment will help achieve a good signal-to-noise ratio and minimize impedance mismatch (Fig. 2.30 a, d–f).

#### 2.7.2.2 In Underwater Studies

An important issue with respect to fishes and invertebrates is their sensitivity to particle motion that accompanies sound transmission, rather than to sound pressure. Particle motion comprises particle displacement, particle velocity, and particle acceleration (ISO 18405 2017[27]) and differs from sound pressure in that it is a vector quantity. In contrast, sound pressure is a scalar quantity, acting in all directions.

Popper and Hawkins (2018) reported that it is commonplace to characterize underwater sound by the sound pressure alone, because it is easily measured by a hydrophone, and then to estimate the particle motion from the sound pressure measurements and the acoustic properties of the medium. This is relatively easy in an acoustic free-field (i.e., no nearby boundaries to sound propagation). However, near acoustic boundaries (like the seabed and the sea surface), the relationship between pressure and particle motion becomes complex and so, particularly in shallow waters that are inhabited by many fishes and invertebrates, measuring particle motion directly is necessary. The result is a dearth of data on particle motion and its importance to, and potential effects upon, animals. Although there are excellent hydrophones for monitoring sound pressure, there are far fewer devices for detecting and analyzing particle motion.

Popper and Hawkins (2018) described the many problems with measuring particle motion in a tank and recommended that measurements be taken in the field, or at least in a specially designed sound exposure chamber to control the relative magnitudes of particle motion and sound pressure. To make particle motion measurements, it is necessary to mount three orthogonally orientated vector sensors together to monitor the three spatial components of particle motion. Any sound can thus be resolved into its directional components and the direction to the sound source may be determined. Calibrated particle motion measurement systems are commercially available, but expensive. An alternative approach is to measure the sound pressure gradient in the water to derive the particle motion in a particular direction.

Many studies have used custom-built particle motion sensors for studying the impacts of anthropogenic activities on fish (e.g., Campbell et al. 2019; Solé et al. 2017; van der Knaap et al. 2021). GeoSpectrum Technologies Inc. offers a few choices for off-the-shelf particle motion sensors in their M20 line of products. Each device consists of an omnidirectional acoustic pressure sensor co-located with three (or two) dipole sensors that measure the amplitude and phase of particle motion in the three (or two) orthogonal directions. Being lightweight and having a small form factor (e.g., the M20–040 has a 64 mm diameter and is 179 mm tall; Fig. 2.31), they are

[27] https://www.iso.org/standard/62406.html; accessed 8 Mar. 2021.

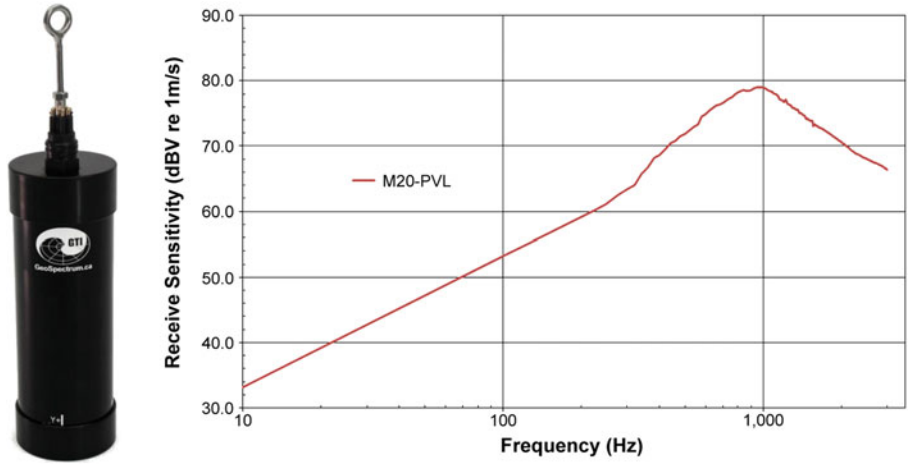



**Fig. 2.31** Photograph (left) and receiving frequency response (right) of GeoSpectrum M20–040. Note that the units of the calibration curve are in terms of particle velocity level (PVL): dBV re 1 m/s. Permission to reprint from GeoSpectrum Technologies Inc. (http://www.GeoSpectrum.ca/; accessed 15 Mar. 2021)

preferred over traditional hydrophone arrays for assessing directionality, especially for use on small unmanned underwater vehicles (e.g., Stinco et al. 2019). The M20 devices support directionality assessments over a frequency range of 1 Hz to 3 kHz, and the bearing uncertainty increases with decreasing frequency and decreasing SNR. Erbe et al. (2017) used a GeoSpectrum M20 to determine sound pressure, particle displacement, particle velocity, and particle acceleration from recreational swimmers, kayakers, and divers.

### 2.7.3 Smartphone Applications

Smartphone applications have put bioacoustic research in the hands of hobbyists and citizen scientists. Applications are inexpensive, rapidly evolving, and available on both Android based phones and iPhones. These applications are well-suited for classroom and field demonstrations of bioacoustic research. The microphone and soundcard in cellphones from different manufacturers determine the frequency range and level of the sounds recorded and the type of analysis possible. A researcher needs to know the frequency range and amplitude sensitivity of the cellphone to ensure that the sounds of the target animals can be appropriately captured. Applications used in battery-operated cellphones provide the ability to select a recording time and duration for long-term, remote monitoring of ambient and animal sounds.

## 2.8 Summary

Technology used in bioacoustic research is changing rapidly. This chapter describes currently used equipment in bioacoustic studies, along with references and websites. The chapter starts with an introduction to the nomenclature used in the industry, describing these as they apply to animal bioacoustic research. An understanding of the terminology would assist a bioacoustician with choosing appropriate equipment with characteristics suitable for a particular study. Instruments that form a complete recording or playback setup are described in light of these characteristics, along with mentions of a few of the commonly used products available in the market. Considerations such as electronic noise, aliasing, sensitivity, resolution, and dynamic range are discussed for both terrestrial and underwater equipment. Autonomous recorders, that offer pre-packaged programmable solutions for passive acoustic monitoring, are also discussed. The discussions cover several indicative bioacoustic studies (targeting a wide variety of fauna) that highlight the use of specific equipment

for different purposes and under different conditions. Other related types of equipment used in closely related fields (such as biotremology, particle velocity measurement, etc.) are highlighted.

A priori knowledge of the target animal's sounds is helpful in selecting appropriate equipment. Sensing and recording equipment needs to be appropriate for the environmental conditions being studied. This chapter summarizes how to select and operate microphones and hydrophones, digital recorders, automated recording systems, amplifiers, filters, sound pressure level meters, and cellphone applications. Knowing the equipment specifications and selecting components to match in frequency range and amplitude sensitivity is important. The dynamic range, amplitude sensitivity, and frequency response of each piece of equipment in a recording setup must match and suit the types of sound (i.e., their level and frequency range) intended to be recorded. Periodic calibrations of microphones and hydrophones are necessary to ensure accurate measurements are made, and the methods are described herein. With their wide availability and ease of use, smartphone driven approaches are gaining popularity lately. The chapter aims to offer the reader a firm grounding with the concepts and available equipment options in bioacoustics. Pointers to seek further understanding are provided along with information about online resources that could offer more up-to-date information on the topic.

## 2.9 Additional Resources

Information about recording equipment:

- Review by the Macaulay Library of the Cornell Laboratory of Ornithology: https://www.macaulaylibrary.org/resources/audio-recording-gear/; accessed 30 Jan. 2021.
- Introductory guide on instruments and techniques for bioacoustics by the Interdisciplinary Center for Bioacoustics and Environmental Research, University of Pavia: http://www.unipv.it/cibra/edu_equipment_uk.html; accessed 30 Jan. 2021.
- Marco Pesente's blog on getting started with nature recording: http://www.naturesound.it/; accessed 6 Sep. 2021.
- Useful instructions on how to build your own DIY microphones can be found on the email discussion lists *naturerecordists* (naturerecordists@yahoogroups.com) and *micbuilders* (micbuilders@yahoogroups.com).
- For biotremology, recent reviews that discuss sensor possibilities as well as playback equipment include Wood and O'Connell-Rodwell (2010) and Elias and Mason (2014). For a thorough discussion of considerations for vibrational playback experiments, we suggest Cocroft et al. (2014b). An email discussion list of vibrational communication researchers can be found at biotremology@googlegroups.com.

Smartphone applications:

- How to record birds for fun and science and with a cellphone: https://www.allaboutbirds.org/news/how-to-record-bird-sounds-with-your-smartphone-our-tips/; accessed 30 Jan. 2021.

**Acknowledgments** SM thanks Holger Klinck, Director, K. Lisa Yang Center for Conservation Bioacoustics, Cornell Lab of Ornithology, for his support and advice on some of the topics covered in the chapter. Thanks are extended by LAM to Lasse Jakobsen and Magnus Wahlberg, Institute of Biology, University of Southern Denmark, Odense, Denmark, and Jakob Tougaard and Peter T. Madsen, Institute for Bioscience, Aarhus University, Aarhus, Denmark, for comments on this chapter. WLG thanks Natalie Gannon and Mithriel for information and photographs on marine acoustic programs in Puerto Rico. Michael O'Farrell provided current notes on Anabat and other bat detector technology. Dean Julie Coonrod, University of New Mexico, provided academic support for completion of this project. GP thanks Marco Pesente for his contribution of material about DIY microphones.

## References

Beason RD, Rüdiger Riesch R, Koricheva J (2018) AURITA: an affordable, autonomous recording device for acoustic monitoring of audible and ultrasonic

frequencies. Bioacoustics 28(4):381–396. https://doi.org/10.1080/09524622.2018.1463293
Bell PD (1980) Transmission of vibrations along plant stems: implications for insect communication. Journal of the New York Entomological Society 88:210–216
Belwood JJ, Morris GK (1987) Bat predation and its influence on calling behavior in neotropical katydids. Science 238:64–67. https://doi.org/10.1126/science.238.4823.64
Brüel and Kjær (1982) Condenser-microphones. Brüel & Kjær, Denmark: 1–146. http://www.bkhome.com/doc/be0089.pdf
Brüel and Kjær (2001) Environmental noise. Brüel & Kjær, Denmark: 1–67. http://www.bkhome.com/doc/br1626.pdf
Buck CL, Malavar JC, George O, Koob GF, Vendruscolo LF (2014) Anticipatory 50 kHz ultrasonic vocalizations are associated with escalated alcohol intake in dependent rats. Behav Brain Res 271:171–176
Buzzetti F, Brizio C, Pavan G (2020) Beyond the audible: wide band (0–125 kHz) field investigation on Italian Orthoptera (Insecta) songs. Biodiversity Journal 11(2): 443–496
Campbell J, Sabet SS, Slabbekoorn H (2019) Particle motion and sound pressure in fish tanks: a behavioural exploration of acoustic sensitivity in the zebrafish. Behavioural Processes 164:38–47
Caruso F, Sciacca V, Bellia G, De Domenico E, Larosa G, Papale E, Pellegrino C, Pulvirenti S, Riccobene G, Simeone F, Speziale F, Viola S, Pavan G (2015) Size distribution of sperm whales acoustically identified during long term deep-sea monitoring in the Ionian Sea. PLoS One 10(12):e.0144503. https://doi.org/10.1371/journal.pone.0144503
Cocroft RB (1996) Insect vibrational defence signals. Nature 382:679–680
Cocroft RB, Rodríguez RL (2005) The behavioral ecology of insect vibrational communication. BioScience 55: 323. https://doi.org/10.1641/0006-3568(2005)055[0323:TBEOIV]2.0.CO;2
Cocroft RB, Gogala M, Hill PSM, Wessel A (eds) (2014a) Studying vibrational communication. Springer-Verlag, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-43607-3
Cocroft RB, Hamel JA, Su Q, Gibson J (2014b) Vibrational playback experiments: challenges and solutions. In: Janik EM, McGregor P (eds) Studying vibrational communication. Springer, Heidelberg, pp 249–274
Čokl A, Prešern J, Virant-Doberlet M, Bagwell GJ, Millar JG (2004) Vibratory signals of the harlequin bug and their transmission through plants. Physiological Entomology 29:372–380. https://doi.org/10.1111/j.0307-6962.2004.00395.x
Čokl A, Zorović M, Žunič A, Virant-Doberlet M (2005) Tuning of host plants with vibratory songs of *Nezara viridula* L (Heteroptera: Pentatomidae). J Exp Biol 208:1481–1488. https://doi.org/10.1242/jeb.01557
Čokl A, Zorović M, Kosi AŽ, Stritih N, Virant-Doberlet M (2014) Communication through plants in a narrow frequency window. In: Janik EM, McGregor P (eds) Studying vibrational communication. Springer, Berlin, pp 171–195. https://doi.org/10.1007/978-3-662-43607-3_10
Collins J, Jones G (2009) Differences in bat activity in relation to bat detector height: implications for bat surveys at proposed windfarm sites. Acta Chiropterologica 11:343–350
Dannhof BJ, Bruns V (1991) The organ of Corti in the bat *Hipposideros bicolor*. Hear Res 53(2):253–268
De Souza LR, Kasumovic MM, Judge KA (2011) Communicating male size by tremulatory vibration in a Columbian rainforest katydid, *Gnathoclita sodalis* (Orthoptera, Tettigoniidae). Behaviour 148:341–357. https://doi.org/10.1163/000579511X559418
Elias DO, Mason AC (2010) Signaling in variable environments: substrate-borne signaling mechanisms and communication behavior in spiders. In: O'Connell-Rodwell CE (ed) The use of vibrations in communication: properties, mechanisms and function across taxa. Research Signpost, Kerala
Elias DO, Mason AC (2014) The role of wave and substrate heterogeneity in vibratory communication: practical issues in studying the effect of vibratory environments in communication. In: Janik EM, McGregor P (eds) Studying vibrational communication. Springer, Heidelberg, pp 215–247. https://doi.org/10.1007/978-3-662-43607-3_12
Erbe C (2009) Underwater noise from pile driving in Moreton Bay, Qld. Acoustics Australia 37(3):87–92
Erbe C, Verma A, McCauley R, Gavrilov A, Parnum I (2015) The marine soundscape of the Perth Canyon. Progress in Oceanography 137:38–51. https://doi.org/10.1016/j.pocean.2015.05.015
Erbe C, Parsons M, Duncan AJ, Lucke K, Gavrilov A, Allen K (2017) Underwater particle motion (acceleration, velocity and displacement) from recreational swimmers, divers, surfers and kayakers. Acoust Aust 45:293–299. https://doi.org/10.1007/s40857-017-0107-6
Erbe, C., Marley, S., Schoeman, R., Smith, J. N., Trigg, L., and Embling, C. B. (2019). The effects of ship noise on marine mammals--A review. Front Mar Sci 6, 606. https://doi.org/10.3389/fmars.2019.00606.
Favali P, Chierici F, Marinaro G, Giovanetti G, Azzarone A, Beranzoli L, De Santis A, Embriaco D, Monna S, Lo Bue N, Sgroi T, Cianchini G, Badiali L, Qamili E, De Caro MG, Falcone G, Montuori C, Frugoni F, Riccobene G, Sedita M, Barbagallo G, Cacopardo G, Calì C, Cocimano R, Coniglione R, Costa M, D'Amico A, Del Tevere F, Distefano C, Ferrera F, Valentina Giordano V, Massimo Imbesi M, Dario Lattuada D, Migneco X, Musumeci M, Orlando A, Papaleo R, Piattelli P, Raia G, Rovelli X, Sapienza P, Speziale F, Trovato A, Viola S, Ameli F, Bonori M, Capone A, Masullo R, Simeone F, Pignagnoli L, Zitellini N, Bruni F, Gasparoni F,

Pavan G (2013) NEMO-SN1 Abyssal Cabled observatory in the Western Ionian Sea. IEEE J Ocean Engineer 38(2):358–374

Gannon WL, Lawlor TL (1989) Variation in the chip vocalization of three species of Townsend chipmunks (genus *Eutamias*). J Mammal 70:740–753

Gorresen PM, Miles AC, Todd CM, Bonaccorso FJ, Weller TJ (2008) Assessing bat detectability and occupancy with multiple automated echolocation detectors. J Mamm 89:11–17

Hamel JA, Cocroft RB (2019) Maternal vibrational signals reduce the risk of attracting eavesdropping predators. Front Ecol Evol 7:614. https://doi.org/10.3389/fevo.2019.00204

Hill PSM (2008) Vibrational communication in animals. Harvard University Press, Cambridge, MA

Hill AP, Prince P, Covarrubias EP, Doncaster CP, Snaddon JL, Rogers A (2018) AudioMoth: Evaluation of a smart open acoustic device for monitoring biodiversity and the environment. Meth Ecol Evol:1–13

Hill P, Lakes-Harlan R, Mazzoni V, Narins PM, Virant-Doberlet M, and Wessel A. eds. (2019). Biotremology: studying vibrational behavior. Springer International Publishing: Cham. https://doi.org/10.1007/978-3-030-22293-2

Jensen ME, Miller LA (1999) Echolocation signals of the bat, Eptesicus serotinus, recorded using a vertical microphone array: effect of flight altitude on searching signals. Behav Ecol Sociobiol 47:60–69

Klinck H, Winiarski D, Mack RC, Tessaglia-Hymes CT, Ponirakis DW, Dugan PJ, Jones C, Matsumoto H (2020) The ROCKHOPPER: a compact and extensible marine autonomous passive acoustic recording system. In: OCEANS 2020 MTS/IEEE Global. IEEE, pp 1–7

Lammers MO, Brainard RE, Au WW, Mooney TA, Wong KB (2008) An ecological acoustic recorder (EAR) for long-term monitoring of biological and anthropogenic sounds on coral reefs and other marine habitats. J Acoust Soc Am 123(3):1720–1728

Lynch E, Joyce D, Fristrup K (2011) An assessment of noise audibility and sound levels in U.S. National Parks. Landscape Ecol 26:1297. https://doi.org/10.1007/s10980-011-9643-x

Magal C, Schöller M, Tautz J, Casas J (2000) The role of leaf structure in vibration propagation. J Acoust Soc Am 108:2412–2418. https://doi.org/10.1121/1.1286098

McCauley RD, Thomas F, Parsons MJG, Erbe C, Cato D, Duncan AJ, Gavrilov AN, Parnum IM, Salgado-Kent C (2017) Developing an underwater sound recorder. Acoust Aust 45(2):301–311. https://doi.org/10.1007/s40857-017-0113-8

McNett GD, Cocroft RB (2008) Host shifts favor vibrational signal divergence in *Enchenopa binotata* treehoppers. Behav Ecol 19:650–656. https://doi.org/10.1093/beheco/arn017

Mennill DJ, Battiston M, Wilson DR, Foote JR, Doucet SM (2012) Field test of an affordable, portable, wireless microphone array for spatial monitoring of animal ecology and behaviour. Methods Ecol Evol 3(4):704–712

Michelsen A, Fink F, Gogala M, Traue D (1982) Plants as transmission channels for insect vibrational songs. Behav Ecol Sociobiol 11:269–281

Miller BS, Barlow J, Calderan S, Collins K, Leaper R, Olson P, Ensor P, Peel D, Donnelly D, Andrews-Goff V, Olavarria C, Owen K, Rekdahl M, Schmitt N, Wadley V, Gedamke J, Gales N, Double MC (2015) Validating the reliability of passive acoustic localisation: a novel method for encountering rare and remote Antarctic blue whales. Endang Species Res 26:257–269

Moir HM, Jackson JC, Windmill JFC (2013) Extremely high frequency sensitivity in a 'simple' ear. Biol Lett 9:20130241

Morris GK (1980) Calling display and mating behaviour of *Copiphora rhinoceros* Pictet (Orthoptera: Tettigoniidae). Anim Behav 28:42-IN1. https://doi.org/10.1016/S0003-3472(80)80006-6

Morris GK, Mason AC, Wall P, Belwood JJ (1994) High ultrasonic and tremulation signals in neotropical katydids (Orthoptera: Tettigoniidae). J Zool 233:129–163. https://doi.org/10.1111/j.1469-7998.1994.tb05266.x

Mortimer B (2017) Biotremology: do physical constraints limit the propagation of vibrational information? Anim Behav 130:165–174. https://doi.org/10.1016/j.anbehav.2017.06.015

Nosengo N (2009) The Neutrino and the whale. Nature 462:560–561

O'Connell-Rodwell CE (2010) The use of vibrations in communication: properties, mechanisms and function across taxa. Research Signpost, Kerala

Obrist MK, Pavan G, Sueur J, Riede K, Llusia D, Márquez R (2010) Bioacoustic approaches in biodiversity inventories. In: Manual on field recording techniques and protocols for all taxa biodiversity inventories. Abc Taxa 8:68–99. ISSN 1784-1283 (hard copy) ISSN 1784-1291 (online pdf). Available at http://www.abctaxa.be/volumes/volume-8-manual-atbi/Chapter-5/

Pavan G (2017) Fundamentals of soundscape conservation. In: Farina A, Gage SH (eds) Ecoacoustics: the ecological role of sound. Wiley, Hoboken, pp 235–258

Pavan G, Borsani JF (1997) Bioacoustic research on cetaceans in the Mediterranean Sea. Mar Freshw Behav Physiol 30:99–123

Pavan G, Manghi M, Fossati C (2001) Software and hardware sound analysis tools for field work. Proc. 2nd Symposium on Underwater Bio-sonar and Bioacoustic Systems. Proc. I.O.A. 23(part 4):175–183

Pavan G, Fossati C, Caltavuturo G (2013) Marine bioacoustics and computational bioacoustics at the University of Pavia (Italy). In: Adam O, Samaran F (eds) Detection classification and localization of marine mammals using passive acoustics. 2003–2013:

10 years of international research. DIRAC NGO, Paris, pp 3–25. 1–298. ISBN 978-2-7466-6118-9
Pavan G, Favaretto A, Bovelacci B, Scaravelli D, Macchio S, Glotin H (2015) Bioacoustics and ecoacoustics applied to environmental monitoring and management. Rivista Italiana di Acustica 39(2):68–74
Payne K, Langbauer WR Jr, Thomas EM (1986) Infrasonic calls of the Asian Elephant (*Elephas maximus*). Behav Ecol Sociobiol 18:297–301
Poole JH, Payne K, Langbauer WR, Moss CJ (1988) The social context of some very low frequency calls of African elephants. Behav Ecol Sociobiol 22:385–392
Popper AN, Hawkins AD (2018) The importance of particle motion to fishes and invertebrates. J Acoust Soc Am 143(1):470–488. https://doi.org/10.1121/1.5021594
Rajaraman K, Godthi V, Pratap R, Balakrishnan R (2015) A novel acoustic-vibratory multimodal duet. J Exp Biol 218:3042–3050. https://doi.org/10.1242/jeb.122911
Rayburn RA (2011) Eargle's the microphone book: from mono to stereo to surround: a guide to microphone design and application, 3rd edn. Elsevier, Waltham, MA
Righini R, Pavan G (2019) First assessment of the soundscape of the Integral Nature Reserve "Sasso Fratino" in the Central Apennine, Italy. Biodiversity Journal 21(1):4–14. https://doi.org/10.1080/14888386.2019.1696229
Sarria-S FA, Buxton K, Jonsson T, Montealegre ZF (2016) Wing mechanics, vibrational and acoustic communication in a new bush-cricket species of the genus *Copiphora* (Orthoptera: Tettigoniidae) from Colombia. Zoologischer Anzeiger - A Journal of Comparative Zoology 263:55–65. https://doi.org/10.1016/j.jcz.2016.04.008
Sciacca V, Caruso F, Beranzoli L, Chierici F, De Domenico E, Embriaco D, Favali P, Giovanetti G, Larosa G, Marinaro G, Papale E, Pavan G, Pellegrino C, Pulvirenti S, Simeone F, Viola S, Riccobene G (2015) Annual acoustic presence of fin whale (*Balaenoptera physalus*) offshore Eastern Sicily, Central Mediterranean Sea. PLoS One 10(11): e0141838. https://doi.org/10.1371/journal.pone.0141838
Solé M, Sigray P, Lenoir M, Van Der Schaar M, Lalander E, André M (2017) Offshore exposure experiments on cuttlefish indicate received sound pressure and particle motion levels associated with acoustic trauma. Sci Rep 7(1):1–13
Stinco P, Tesei A, Biagini S, Micheli M, Ferrri G, LePage KD (2019, June). Source localization using an acoustic vector sensor hosted on a buoyancy glider. In OCEANS 2019-Marseille (pp. 1–5). IEEE. https://doi.org/10.1109/OCEANSE.2019.8867452
Streicher R, Everest FA (1998) The new stereo sound book. AES Publishing, Pasadena, CA
Thode A, Skinner J, Scott P, Roswell J, Strailey J, Folkert K (2010) Tracking sperm whales with a towed acoustic vector sensor. J Acoust Soc Am 128(5):2681–2694
van der Knaap I, Reubens J, Thomas L, Ainslie MA, Winter HV, Hubert J, Martin B, Slabbekoorn H (2021) Effects of a seismic survey on movement of free-ranging Atlantic cod. Curr Biol 31(7):1555–1562.e4. https://doi.org/10.1016/j.cub.2021.01.050
Virant-Doberlet M, Čokl A (2004) Vibrational communication in insects. Neotropical Entomology 33:121–134. https://doi.org/10.1590/S1519-566X2004000200001
Wahlstrom S (1985) The parabolic reflector. J Audio Eng Soc 33(6):418–429
Weller TJ, Baldwin JA (2012) Using echolocation monitoring to model bat occupancy and inform mitigations at wind energy facilities. J Wildl Manag 76:619–631
Whytock RC, Christie J (2017) Solo: an open source, customizable and inexpensive audio recorder for bioacoustic research. Meth Ecol Evol 8:308–312
Wiggins SM, Hildebrand JA (2007) High-frequency acoustic recording package (HARP) for broad-band, long-term marine mammal monitoring. In: International Symposium on Underwater Technology and Workshop on Scientific Use of Submarine Cables and Related Technologies 2007 (IEEE, Tokyo, Japan): 551–557
Wood JD, O'Connell-Rodwell CE (2010) Studying vibrational communication: equipment options, recording, playback and analysis techniques. In: The use of vibrations in communication: properties, mechanisms and function across taxa. Research Signpost, Kerala, pp 163–182
Zimmer WMX (2011) Passive acoustic monitoring of cetaceans. Cambridge University Press, Cambridge
Zimmer WMX (2013) Range estimation of cetaceans with compact volumetric arrays. J Acoust Soc Am 134(3): 2610–2618
Zotter F, Frank M (2019) Ambisonics. Springer International Publishing, Cham. ISBN 978-3-030-17206-0. https://doi.org/10.1007/978-3-030-17207-7

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



# Collecting, Documenting, and Archiving Bioacoustical Data and Metadata

3

William L. Gannon, Rebecca Dunlop, Anthony Hawkins, and Jeanette A. Thomas

## 3.1 Introduction

Over the last 100 years, bioacoustical research has led to many important discoveries about the role of sounds in animal behavior. Over time, best practices have evolved in bioacoustical research; often through trial and error. In this chapter, these best practices, based on the literature and the co-authors' experiences and opinions, are summarized. We recommend methods to properly collect and conserve data, use appropriate equipment, save time, and perhaps even make a study more affordable. It is advised, of course, that researchers conduct a current literature review before beginning their work, as developments in technique and technology are moving at a fast pace.

Although methods in bioacoustical studies are typically non-invasive, research should be conducted in an ethical way and any necessary permits obtained. Bioacoustical research should be able to be repeated reliably, where another investigator should be able to understand the circumstances of the recordings, replicate and apply the results, and be reassured the methods were appropriate for the goals of the study. Detailed logs of recordings are important and should include names of researchers; date and time; location; ambient conditions; equipment specifications; species, age, and sex; and behavioral context of the animal during the recording. Details of data collection and signal analysis should accompany any results, such as frequency range, sampling rate, bit-resolution, analysis bandwidth and interval, amplitude range, and any filtering or weightings used.

Here, we also discuss special considerations, or adaptation of methods, for acoustic studies in aquatic versus terrestrial field environments, as well as considerations for studies on captive animals. The "playback" technique, where a sound is played back to an animal and response noted, is a common method used in bioacoustical studies and this chapter provides recommendations for designing a robust playback study. Finally, methods for data archival, and current repositories for bioacoustical data, are provided as a resource for those interested in examining existing data or preserving their own recordings.

Jeanette A. Thomas (deceased) contributed to this chapter while at the Department of Biological Sciences, Western Illinois University-Quad Cities, Moline, IL, USA

W. L. Gannon (✉)
Department of Biology, Museum of Southwestern Biology and Graduate Studies, University of New Mexico, Albuquerque, NM, USA
e-mail: wgannon@unm.edu

R. Dunlop
School of Biological Sciences, University of Queensland, Brisbane, QLD, Australia
e-mail: r.dunlop@uq.edu.au

A. Hawkins
The Aquatic Noise Trust, Kincraig, Blairs, Aberdeen, UK

© The Author(s) 2022

C. Erbe, J. A. Thomas (eds.), *Exploring Animal Behavior Through Sound: Volume 1*,
https://doi.org/10.1007/978-3-030-97540-1_3

## 3.2 Ethical Research

As with all scientific endeavors, bioacousticians work to answer questions and address hypotheses by observing or manipulating the natural world. There is an ethical obligation to document procedures and methods, so that reported results are understandable and reproducible by other researchers. A reliable way for understanding data, and how they were collected, is by documenting metadata associated with a recording. Metadata are the description of basic information collected at the time of the recording, such as the recordist; date and time; specific location (GPS coordinates); equipment and settings; water depth or altitude; water or air medium; water or air temperature/humidity; weather conditions; and species, sex, age, and behavior of the animals. Knowing the who, what, when, and where, of acoustic recordings makes acoustic data more useful and allows a review of methods by other researchers to validate or supplement data.

Although bioacoustical studies are usually non-invasive, investigators need to consider and minimize any potential effects of their work on animals (e.g., avoid playbacks of extremely loud or injurious sounds that could disturb animals in critical breeding and feeding areas). In many cases, animal ethics permits and/or research permits are needed from the country, state, county, or any other political entity in which the study will be conducted. If the species is endangered, additional permits may be required. Most research institutions receiving funding from the USA government require investigators to submit an animal research protocol to an Institutional Animal Care and Use Committee (IACUC) for approval before conducting research involving any animals. Ethical conduct of research goes beyond satisfying the requirements of the IACUC and includes responsible data collection and management, appropriate statistical analyses, thorough presentation and archival of data, and a study that is reproducible. Additionally, research should be reported, peer-reviewed, and published ethically. This falls under research ethics principles and studies that are conducted with scientific integrity (Fig. 3.1). Most researchers consider their work with animals to be harmless and therefore ethical. However, the process of thinking through how animals could be affected, and proposing research methods during the preparation of an IACUC protocol can be very instructive. In some cases, preparing a protocol for review can save a project from mistakes (such as low statistical power, inadequate or illegal animal housing or handling methods, unnecessary duplication, unnecessary expense, or unrecognized alternative hypotheses). In fact, developing a research protocol can serve to make the research more robust.

Gannon (2014) provided two examples that illustrate a potentially unethical study and posed the question of whether a research permit was needed. In 1991, a rare migrant yellow-green vireo (*Vireo flavoviridis*) was spotted at protected parklands in Rattlesnake Springs, New Mexico, USA. The sighting was announced on the rare-bird hotline and a number of people went to the area to view the bird and to add it to their "life list." During this time, a PhD student was collecting goldfinches (*Spinus tristis*). Knowing that genetic material and voucher specimens are important to taxonomic and conservation research, he decided to collect the rare bird for a museum research collection. To entice the bird to an unprotected area for easy and legal collection, he recorded calls of the vireo and then played them back where he could legally collect the bird. The birding community became incredulous and angry. Was it ethical to record and use playbacks of this species' calls to lure the bird to an unprotected area for collection (see Gluck 1998)?

More recently, as characterized in Fig. 3.2, a smartphone birding application was used to lure a male common yellowthroat (*Geothlypis trichas)* into view. White (2013) described that broadcasting calls, using a smartphone application, generally elicits a quick response from a normally concealed bird. Possibly thinking the sounds were from another male of his species and threatening his territory, the male yellowthroat swooped down right in front of a birding tour




**Fig. 3.1** A collage of common reference materials and journals that are used to advise on the responsible conduct of research with animals. Considerations of the integrity of the scientific process and the ethics of how a study is conducted undoubtedly produce better science

and was photographed. Is it ethical to lure a bird to impress a tour group or does the playback burden the bird with unnecessary stress, perhaps reducing his fitness? Should acoustic luring be prohibited for all bird species or for only endangered animals? Conversely, should these techniques be encouraged in order to raise awareness of wild things to a public who are increasingly alienated from nature?

Ethical treatment of animals serves to make a research project rigorous and results stronger. Given the personnel time to design experiments, obtain permits, and conduct bioacoustical research, and given the expense and potential disturbance to animals, is the project worth doing? If it is worth doing, it is worth doing well.

## 3.3 Good Practices in Bioacoustical Studies

Once research questions have been developed and equipment has been selected (see Chap. 2 on equipment choices), recording can begin! Animals can be recorded in a controlled laboratory or in the field. Bioacousticians often need to be innovative when collecting acoustic data in field situations because additional equipment, AC-power, and access to repairs are not always available. Below is a summary of some recommendations for beginning bioacousticians. All suggestions are relevant to both terrestrial and aquatic environments unless identified otherwise.




**Fig. 3.2** Caricature of an ornithologist luring a bird by playback of bird calls (with permission of the illustrator Rohan Chakravarty)

### 3.3.1 Recording Sounds

It is best to work toward making the cleanest recording possible for accurate acoustic analysis. Be sure that you have a solid understanding of the gain and level controls on your recorder. The gain and level meter work in concert and the person making the recording needs to be comfortable with these settings before serious acoustic research begins. Ideally the entire recording chain should be calibrated. Calibration generally refers to correlating the readings of an instrument with those of a standard for the purpose of checking the instrument's accuracy. When recording sound, a calibration signal (a pure tone) of known frequency and amplitude should be placed at the beginning of all recordings. Some recorders have a built-in calibration tone. The tone also can be used to mark an important section of the recording. Having a calibration tone on a recording allows measurement of absolute amplitude, rather than just relative amplitude. This step is necessary if the researcher wants to report source-levels of animal or environmental sounds. Calibrating recording equipment is referred to in Chap. 2 of this volume. Ideally the distance to the sound source (vocalizing animals in our case) should be known. A common "trick" is dropping a colored poker chip at the point where the recording is started and then as moving toward the sound source, dropping additional chips until the point where the animal who had been calling has presumable run off. The distance can then easily be measured between chips. Absolute distance and calibration of the recording system is difficult in field studies.

If more than one channel is available on a recorder, use one channel to narrate metadata and the animals' behaviors with the second channel dedicated to recording animal sounds. This allows all details and conditions of the situation to be documented in real-time and synchronized with the animals' sounds and behaviors. After each session, the researcher should listen to the recordings to make sure signals were recorded and the equipment was working properly. We recommend making a copy of each recording and storing the backup and the original in different places.

When possible, use battery-power or direct-current (DC), rather than alternating-current (AC) wall- or shore-power. Using batteries eliminates background electronic noise and provides portability of the equipment. AC-power can create a 50-Hz (European power) or 60-Hz (North American power) hum or background noise on a recording. This frequency-specific noise is easy to recognize and filter-out, preferably during the recording. However, if the animal produces low-frequency signals (e.g., 20-Hz calls from some baleen whales, low-frequency knocks and grunts from fish, rumbles by elephants) the recordings should not be filtered. Note that in extremely cold locations, battery-life will be shorter and any type of mechanical components such as belts, gears, toggles, reels, or digital equipment can cease to operate correctly. We recommend that backup batteries be available or on-charge for quick battery exchange.

### 3.3.2 Environmental Conditions

Equipment should be selected based on environmental conditions at the field site including ambient temperature and humidity, prevalence of wind and waves, amount and type of precipitation, and frequency and amplitude of the target species (Fig. 3.3; see Chap. 2 on equipment choices). Before commencing field work, check the weather forecast. Recording animal sounds during precipitation, high wind, or a high sea-state often is futile because incoming signals will be masked. In addition, animals sometimes do not call during these conditions. In terrestrial environments, noise from wind, weather, moving vegetation, or other animal sounds can mask recordings of the target species (see Chap. 5 on the source-path-receiver model for airborne sound). In aquatic habitats, wind, sea-state, breaking waves, precipitation, and other animal sounds can create a noisy background. In both terrestrial and aquatic environments, anthropogenic noise (from vehicles and vessels, industrial operations, military activities, etc.) essentially is omnipresent (see Chap. 7 on soundscapes). If using a remote recording system, protect the unit from the weather and secure it as best possible. Be aware that even in remote locations, theft of field equipment occurs.



**Fig. 3.3** Conditions in the field often contrast sharply from those in a controlled laboratory environment. Working to exclude bats (Townsend's big-eared bat, *Corynorhinus townsendii*) from gold mining operations in Nevada, USA (top left). Recording assures animals are excluded prior to destroying the tunnel system for mineral extraction. Mitigation sites are identified (top right) which are gated and protected for bats to inhabit safely. Occasional sampling is completed by live-capture (bottom left) and acoustic monitoring (bottom right). All photos by authors except bottom left (MNH field biologists collect bat specimens, by Florante A. Cruz; https://www.wikiwand.com/en/UPLB_Museum_of_Natural_History; licensed under CC BY-SA 4.0; https://creativecommons.org/licenses/by-sa/4.0/



**Fig. 3.4** Photographs of researchers in Antarctica recording a killer whale (*Orcinus orca*; left) and Weddell seal (*Leptonychotes weddellii*; right). Equipment is both protected from being molested by the animal but also not prominent so as to not draw the subject's attention. Note the researcher on the right maintains a distance from the seal so as not to disturb it

Documenting the ambient temperature and humidity is especially important when studying ectothermic terrestrial animals, such as reptiles, frogs, toads, insects, or other invertebrates. At low ambient temperatures, ectothermic animals are less active and sounds are lower in frequency than during higher ambient temperatures. For example, studies by Kissner et al. (1997) demonstrated that sounds from ectothermic animals, such as rattlesnakes (*Crotalus viridis*), change with ambient temperature and humidity.

### 3.3.3 Animal Considerations

The transducer should be positioned so target-animal sounds are recorded but the animal does not damage the equipment. An aggressive or curious animal can quickly demolish a recording system (Fig. 3.4). Equipment used in playback studies can be particularly susceptible to an animal attack. The goal of recording is to document sounds from natural circumstances and not from a charging or frightened animal. Captive animals often are curious about a hydrophone or a microphone in their enclosure and can need time to habituate to equipment before undisturbed sounds are produced. Placing the transducer in a protected area or in a protective mesh cage may be necessary.

Researchers should not disturb animals while recording (Fig. 3.5). If possible, the recordist should hide in a blind spot or use an automated recording system with no observer present. Note that sometimes narrating observations of the animal's behavior during the recording is useful which means that the researcher should decide between using a remote setup and a setup where they are nearby. To concurrently monitor animal behavior, a video camera on a tripod can be used, with minimal disturbance to the animal. However, the researcher should be aware that the audio track of a video camera has a limited frequency response and an auto-adaptive level control, meaning these sound recordings should not be relied upon for acoustical analysis. Closed Circuit Television (CCTV), synchronized with omnidirectional microphones on an ultrasonic detector, and coordinated using a mobile phone and speaking clock, has been used to document new vocalizations and activities patterns for barbastelle bats (*Barbastella barbastellus*; Young et al. 2018). With a little ingenuity, a researcher can create a robust recording system.

To save time and expense, it is important to know whether a species has a preferred time of day or season for producing sounds. Many species are most vocal during the breeding season. Some birds and amphibians are most soniferous at dawn and dusk whereas many chorusing

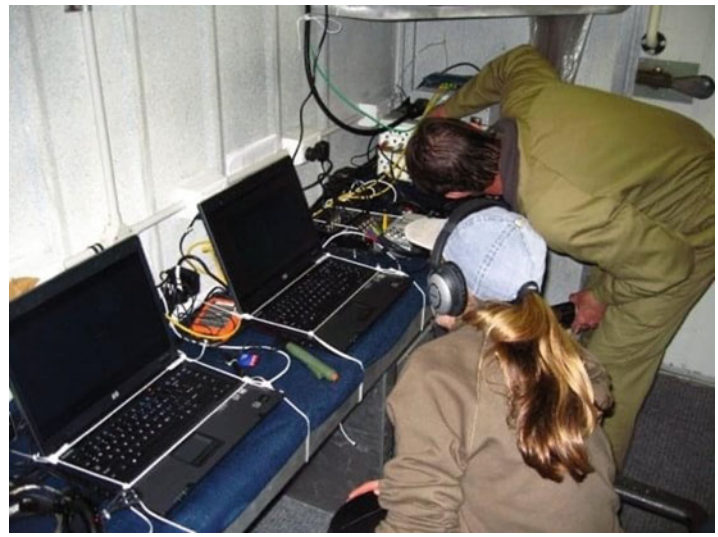


**Fig. 3.5** What could go wrong? In the field, equipment failure is certain. Over-planning, backups, duplicate systems, checklists, and more will help avoid data collection failures

insects primarily produce sounds at dusk. For example, Thomas and DeMaster (1982) showed that Antarctic crabeater seals (*Lobodon carcinophaga*) preferred to call under water between 2100 h and 0500 h and were hauled-out on the ice at other times. If the number of vocalizations was used as a population count, a census of crabeater seals at 1200 h would have yielded a much lower population estimate than a census at 2400 h. Bats, obviously, are active at night. However, there is usually a notable peak of activity approximately 30 minutes after dusk (Kunz and Parsons 2009). Some species (many in the genus *Myotis* and *Tadarida*) are more likely to be recorded during the first four hours of night, while others emerge past midnight (*Euderma*, *Artibeus*). Some bats have multimodal activity patterns (Sherwin et al. 2000) and many sciurids (e.g., *Marmota* and *Neotamias*) actively vocalize in the morning and then again in late afternoon (Gannon 1999). Some species (e.g., prairie dogs, *Cynomys* and pikas, *Ochotona*) are seasonally soniferous all day (Slobodchikoff et al. 1998; Smith et al. 2016).

It is important to know the effects of both time of day and month to interpret the behavioral context of a recording. For example, breeding data from the North American male rufous-sided towhee (*Pipilo erythrophthalmus*) showed that males reached breeding condition around mid-April. Testes were in regression by 20 July and had become inactive by mid- to late-September (Davis 1958). So, if a researcher desires to record sounds of this species associated with breeding, the study should be conducted from mid-April to mid-July. In addition, this species shifts their song to an earlier start time in relation to civil twilight. As day length increases between the spring equinox and the summer solstice, civil twilight occurs earlier in relation to sunrise, causing the dawn calling period to lengthen.

### 3.3.4 Documentation and Data Sheets

Documentation is very important. A logbook should accompany each recording to provide metadata on the recordist; the recording system and equipment settings (e.g., any filter or gain settings); the location, date and time; environmental conditions; types of sounds recorded; the animals' behavior (e.g., breeding, feeding, or socializing); a specific animal number (if marked); and any other circumstances which could be valuable for analysis.

Many devices may record some of the metadata automatically. For instance, the Echo Meter Touch 2 PRO Ultrasonic Module using

**Table 3.1** Sample logbook showing important metadata to be noted. Examples from author (JAT) notes for Weddell seal (*Leptonychotes weddellii*) and sea otter (*Enhydra lutris*)

| Tape | Counter | Collector | Date | Time | Location | Subject | Quality | Comments |
|---|---|---|---|---|---|---|---|---|
| 2 | 234 | JA Thomas | 23 March 2004 | 16:00 | McMurdo | Weddell seal | Poor | Underwater, adult male, 839W, wind 20 knts |
| 13 | 22 | CM Smith | 18 Sept 2004 | 13:15 | Valdez, AK | Sea otter | Excellent | Airborne, mother and pup, unmarked, no wind |

Kaleidoscope Pro software[1] (Wildlife Acoustics, Maynard, MA, USA) records calls to an iPhone or other device and collects metadata about each recording. Metadata can then be displayed with Kaleidoscope software or exported to a spreadsheet. Recording directly to a computer allows time-stamped (and often GPS-stamped) files.

If a datasheet (spreadsheet) is used, put metadata headers as the first column and fill the rows with your observations (Table 3.1). Each sound or bout of sounds should be assigned a unique number for easy reference later, and a variety of variables can then be noted for each sound (Table 3.2). Spreadsheets can be imported directly into a variety of statistical and graphing software products for analyses (see Chap. 9 on analytical approaches). Note that datasheets for playback studies usually include additional variables on animal behavior (Table 3.3).

### 3.3.5 Trouble-shooting Equipment Problems

Often field work is conducted in remote locations, sometimes without easy access to the Internet, electricity, or equipment repairs. Consider all possible equipment problems and always have backups—of everything. A good motto for field work is to "bring one to use and one to lose" (Fig. 3.5). Studies usually are costly and time-consuming—in particular in remote locations. There is nothing worse than a missed field opportunity caused by the lack of a cable or battery.

Bring proper tools to the field site to make repairs: soldering iron, solder, electrical wire, heat-shrink tubing, electrical ties, electrical tape, extra cables and connectors, batteries (preferably rechargeable, with charger), multi-meter, etc. If possible, pack replacement equipment: anemometer, thermometer, laptop with extra charger, external speakers, software for data entry, backup hydrophone or microphone, headset, walkie-talkie, smartphone, microphone for narration onto a PC, and data storage devices (SD-cards, thumb-drive, external hard-drive). Why are duplicates necessary? If you cannot repair something, then use backups so the research effort is not wasted.

Moving or shipping equipment often creates problems with loose connections or fittings. If equipment is not operating properly, tighten fasteners on the equipment housing, make sure circuit boards are seated properly, check that batteries are fully charged, and make sure all cables are connected and working. To check for cable malfunction, use an ohm-meter to make sure the resistance of a cable is zero. If new equipment is used in a study, always unpack it and check its operation in the laboratory before going to the field. Bring manuals for all equipment to the field site or know where to reliably access them.

## 3.4 Playback Methods and Controls

Projections of sounds to animals (or playbacks) are common methods of study in bioacoustics (Fig. 3.6). Several authors have used playbacks to determine the function of a specific animal sound by measuring the animal's behavioral response (Morton and Morton 1998).

[1] https://www.wildlifeacoustics.com/products/echo-meter-touch-2-pro-ios and https://www.wildlifeacoustics.com/products/kaleidoscope-pro; accessed 13 June 2022

**Table 3.2** Sample data sheet for acoustic measurements on airborne vocalizations of California sea lions (*Zalophus californianus*); na not applicable. Courtesy of Schwalm (2012)

| Date | Sound | Dominant start frequency (Hz) | Dominant end frequency (Hz) | Dominant maximum frequency (Hz) | Dominant minimum frequency (Hz) | Harmonics | First harmonic interval (Hz) | First component duration (s) | Interval to second component (s) | Total duration (s) | Rate |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 3 July 2011 | 1 | 642 | 755 | 943 | 491 | Yes | 453 | 0.838 | na | 0.838 | Single |
| 4 July 2011 | 2 | 566 | 566 | 717 | 415 | Yes | 377 | 0.253 | 0.148 | 5.211 | Even, series of 2 |
| 5 July 2011 | 3 | 640 | 534 | 720 | 294 | No | na | 0.139 | na | 0.139 | Single |
| 9 July 2011 | 4 | 614 | 800 | 881 | 480 | Yes | 26 | 0.388 | 0.146 | 3.477 | Accelerate, series of 4 |
| 9 July 2011 | 5 | 587 | 667 | 747 | 427 | Yes | 400 | 0.165 | 0.146 | 5.57 | Irregular, series of 6 |

**Table 3.3** Sample data sheet for playback study with white-cheeked gibbons (*Nomascus leucogenys*); na not applicable. Courtesy of Yegge (2012)

| Session | Date | Time | Type of playback | Animal | Quadrant number | Mate in same quad | Number calls | Approach speaker | Move away from speaker | Stationary | Climb | Move | Groom | Out-of-sight |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3 | 29 Aug | 16:00 | Control | CJ | 2 | Yes | 0 | Yes | No | No | Yes | Run | No | No |
| 3 | 29 Aug | 16:00 | Control | Max | 2 | No | 0 | No | No | Yes | No | na | Yes | No |
| 4 | 30 Aug | 16:30 | Own duet | CJ | 3 | Yes | 2 | No | Yes | No | No | Walk | No | No |
| 4 | 30 Aug | 16:30 | Own duet | Max | 4 | Yes | 4 | No | No | Yes | No | na | Yes | No |



**Fig. 3.6** Playback studies are those by which an animal or group of animals is played their calls (or calls of their conspecifics) back to them and then their response is recorded. Research using playbacks has been used commonly in mammals (such as squirrels, prairie dogs, pika, carnivores, and primates), birds, reptiles, fish, and many others. Painting "His Master's Voice" by Francis Barraud (1856–1924). Source: Victor Talking Machine Company. Public domain; https://commons.wikimedia.org/wiki/File:His_Master%27s_Voice.jpg

Playback studies on fish have been used to determine species recognition from a particular sound, to classify different call types, to identify effects of sound on fish behavior, to study how a call was coded, and to measure acoustic parameters of the call relevant to communication (Zelick et al. 1999). For example, Myrberg and Riggio (1985), studying bicolor damselfish (*Stegastes partitus*), found that males produced sounds more often in response to playbacks of conspecific sounds than to sounds of other species', and responded more readily to sounds from non-resident fish than sounds from their nearest neighbor. Playbacks of male Lake Malawi cichlid fish (*Pseudotropheus zebra*) sounds to female cichlids caused them to lay eggs earlier than control female fish of another Lake Malawi cichlid species (*Pseudotropheus emmiltos*; Amorim et al. 2008). Simpson et al. (2011) played-back ambient sounds of different reefs to coral reef fish and showed that fish approached the sounds of their native coral reef versus sounds from a foreign reef. Hawkins et al. (2014) played back recordings of impulsive pile driving sounds attracting European sprat (*Sprattus sprattus*) in mid-water in the sea (Fig. 3.7).

Many birds respond to playbacks of their own or other animal sounds by approaching the projector and sometimes even attacking the speaker (Fig. 3.8). Emlen (1972) investigated how information is encoded in bird song by altering components of Indigo bunting (*Passerina cyanea*) song and playing-back the modified songs to male territory holders. He quantified the intensity of responses to modified songs and thus inferred the importance of temporal, structural, and syntactical features for both individual- and species-recognition.

Beecher and Burt (2004) played-back territorial sounds from male song sparrows (*Melospiza melodia*) that were in neighboring territories versus distant territories. The males were slower and less likely to fly over and explore the sounds from a neighbor than calls from a distant male. When a song from a distant territorial male was played, the subject almost always matched or replicated the song and approached the speaker as if looking for an intruder. In contrast, when the song of a

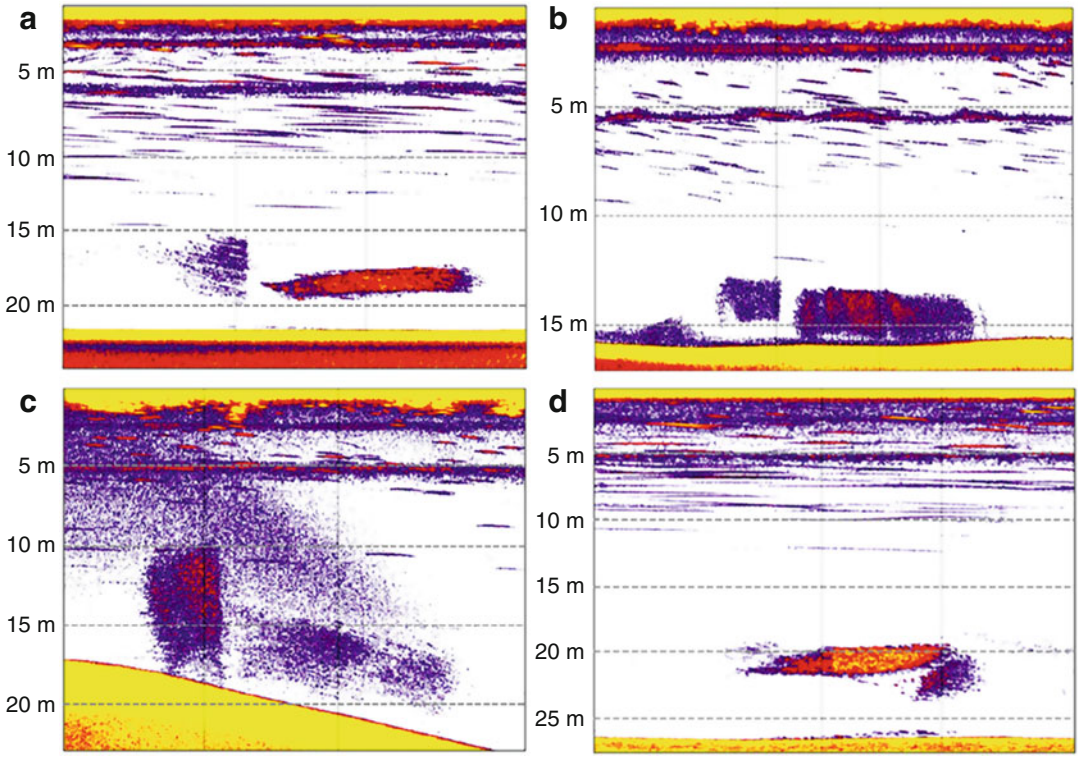



**Fig. 3.7** Responses of sprat (*Sprattus sprattus*) schools to sound exposure. Vertical lines indicate the beginning and end of each sound sequence. (**a**) Echogram of a medium-sized sprat school, cut off abruptly after the beginning of the sound, and reappearing a few seconds later as a denser school slightly closer to the seabed. (**b**) A medium-sized sprat school cut off at the onset of the sound and reappearing seconds later slightly closer to the seabed. (**c**) A large sprat school cut off at the onset of the sound and reappearing at a greater depth at lower density. (**d**) A small sprat school increasing in density in response to sound exposure. From Hawkins et al. 2014. © Acoustical Society of America, 2014. All rights reserved

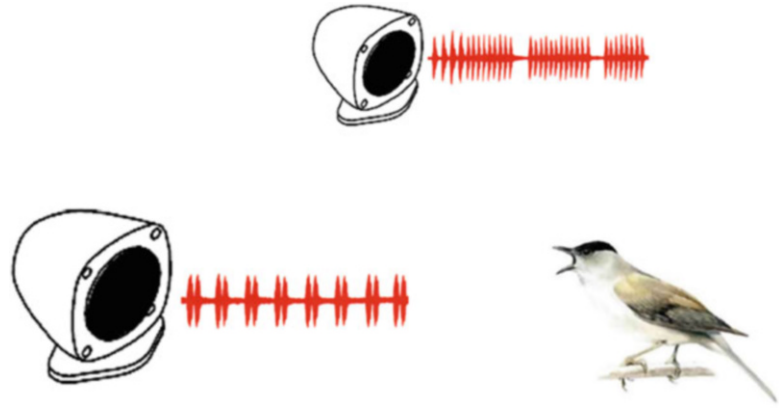


**Fig. 3.8** Diagram of a playback experiment with two different bird songs. The recording and the speakers should match the frequency range and levels of the original signals. Courtesy of G Pavan

neighbor male was played, 85% of the time the subject sang a different song, but one familiar to the neighbor. By responding with a different, but shared song, the subject sparrow indicated it recognized that the sounds were from a neighbor.

Much of the work in determining the function of alarm calls in ground squirrels and prairie dogs (*Spermophilus* and *Cynomys*, respectively) was determined or confirmed by playing-back previously recorded calls to an attentive colony of these rodents in the field and observing their responses (e.g., Slobodchikoff et al. 2009). Prat et al. (2016) used playback techniques of calls recorded from the Egyptian fruit bat (*Rousettus aegyptiacus*) to show that 16 sounds recorded and played-back from this bat provided enough information to identify who was calling, where they were calling from, what they were calling about, and what sort of response the receiver made to the vocalization.

Yegge (2012) and Thomas et al. (2016) reported using playbacks of duets to restore a pair-bond in yellow-cheeked gibbons (*Nomascus gabriellae*). A breeding pair of captive gibbons stopped duetting when construction occurred near their exhibit lasting for about 6 months. Afterwards, the authors played-back sounds of the pair's previous duet, along with a silent- and music-controls. The pair slowly resumed their duet, established a pair-bond, and continued to duet, some 5 years later.

Playback experiments with marine mammals are less common due to the logistical challenges of undertaking these experiments at sea. However, there are a few examples. Weddell seals (*Leptonychotes weddellii*) produced geographically different vocal repertoires that has potential for identifying discrete breeding stocks of Antarctic seals (Thomas et al. 1983). Charrier et al. (2013) used playback methods to confirm that bearded seals (*Erignathus barbatus*) recognized vocalizations of their species from different regions. Male harbor seals (*Phoca vitulina*) that are territorial, use roars given by intruding seals to locate and challenge those intruders (Hayes et al. 2004). Deecke (2003) used playbacks to examine whether captive harbor seals could distinguish sounds from killer whales (*Orcinus orca*) that eat seals versus killer whales that eat fish; the seals exhibited fearful responses when sounds by the former were broadcast. Wild killer whales either approached or ignored playbacks of sounds from another killer whale pod, but did not call in response. However, when their own calls were played, most killer whales approached the source and the entire pod started calling in response (Filatova et al. 2011). Clark and Clark (1980) described right whale (*Balaena australis*) behavior from playback experiments where right whales can differentiate between conspecific sounds and other sounds. Playbacks of their own song or social sounds to wild humpback whales (*Megaptera novaeangliae*) resulted in some animals approaching, some charging the source, and others moving away (Mobley et al. 1988; Tyack 1983).

Before a playback session, the researcher should always check the projected sound near the animal to make sure the sound is not distorted and is of sufficient amplitude to mimic the intended sound. Ideally, playback experiments should be carried out on wild animals that are free to move within their natural habitats. Captive animals often are de-sensitized to reoccurring sounds, and confinement within a small space can greatly alter their behaviors and vocalizations. It is especially important to ensure that playback experiments are carried out under appropriate acoustic conditions, where the transmitted sounds are free from distortion, and reflection and reverberation are minimal. This is a particular problem with playback experiments on fish, where sounds can be greatly altered by the acoustic environment, especially in small aquarium tanks (Parvulescu 1964; Grey et al. 2016; Rogers et al. 2016).

Playback studies require controls to ensure the animal is responding to the projected sound and not to the noise/hum of equipment or the novelty of a new sound. Current sound analysis and sound-generation software allows the manipulation of many sound characteristics that could be used as a control. There are several types of controls used by investigators: 1) Merely turn on the equipment to replicate the electronic/background noise. 2) Play the animal's own sound,

but backwards. This projects the same frequency, amplitude, and time relationships of the actual sound, but in a different order. 3) Play the animal's sound at a higher or lower speed. This transforms the projected sound into a different frequency range. 4) Play a call with parts filtered out. 5) Play something totally novel to the animal, such as sounds from another species it has never encountered, music, machinery noise, or human speech. 6) Play sounds typical of the animal's natural environment.

## 3.5 Considerations for Terrestrial Field Studies

If recording on land, from a vehicle (such as during a truck survey for bat sounds), ground-generated noise can be a problem. In fact, Borkin et al. (2019) reported a negative relationship between bat activity and night-time traffic volume on New Zealand highways; when traffic increased, probability of detecting bats decreased. These researchers used stationary automatic bat detectors to avoid their own road noise. Some solutions include: stopping and turning the vehicle off and recording in silence; using a recently paved asphalt track rather than an older and noisier road or a dirt track; and carrying out vehicle transects using electric vehicles. Road surveys are valuable, but reducing non-biotic noise would make these transects even more valuable. Terrestrial recordings can be contaminated with nearby traffic noise. It is therefore advisable to make a sample recording, check it for ambient noise, and select an optimal quiet area.

Air temperature can be a problem. Thomas, Zinnel, and Ferm (1983), when recording Weddell seal breeding colonies, used water-activated chemical heat packs placed next to recording equipment and batteries in an insulated box to keep equipment warm in the Antarctic for 24-hour periods. In extremely warm locations with high humidity, moisture can collect on recorders or microphones. Placing recording equipment inside an insulated box with desiccants can minimize moisture problems. In rain forests, equipment must be totally waterproof. During periods of heavy rain, sounds from animals will either not be heard or masked by the rain.

A common problem in bioacoustical studies in terrestrial environments is the presence of acoustically-active non-target animals. If a non-target species calls in a specific frequency band, their sounds can perhaps be filtered out, but in many cases, this is not possible. Some analysis software allows to define the frequency and amplitude of a target species' calls and automatically identifies only them in a recording. However, in many cases, finding locations and times when only an individual animal is vocalizing provides the best opportunity to make quality recordings.

A good solution for animals such as bats is to use units which are self-contained and weather resistant (see Chap. 2, section on bat detectors). Each unit can include a receiving transducer, storage device, or laptop programmed to record at intervals and can be powered by rechargeable battery packs or solar panels. Data can be recovered daily, weekly, monthly, or even uploaded in the proximity of Wi-Fi for automated data retrieval. Arrays of bat detectors have been used to record ultrasonic calls of bats, as well as to sample the acoustic landscape, estimate biodiversity, and estimate species density (Carles et al. 2007; Sherwin et al. 2000).

## 3.6 Considerations for Aquatic Field Studies

Studies in freshwater are easier on the equipment than in saltwater environments; saltwater's corrosive properties require that underwater equipment be rinsed with freshwater after use and recorders and hydrophones be wiped down to remove saltwater deposited from the air. It is, of course, good practice to wipe down and dry all equipment, whether it was deployed in saltwater, in freshwater, or on land, after use to avoid any rusting or build-up of deposits.

Maintenance and calibration of equipment such as hydrophones has been shown to be important for long-term monitoring studies and data integrity. This includes considerations such as

the pressure rating on the hydrophone and the length of cable that is waterproofed; the longer the cable, the higher the impedance and the greater the signal attenuation. Some plastic-coated cables, if deployed for long periods, are vulnerable to damage by marine organisms, shark bites, and even sea urchins. Polytetrafluoroethylene (PTFE) coated cables are less susceptible to damage of this kind. In addition, acoustic-release mechanisms (to allow equipment to surface) can malfunction when encrusted by marine creatures. In a review of underwater soundscape ecology to monitor habitat health in general, and fish spawning in particular, Lindseth and Lobel (2018) summarized current recording and sampling methods including metrics commonly used in analyses of aquatic acoustic data. They point out that there have been significant technological advances in equipment, especially hydrophones.

In aquatic situations, there can be electronic interference from improper grounding on the vessel, depending on the types of electronic equipment running onboard (e.g., lights, radios, freezers, generators, winches, fans, air conditioners, or furnaces). A quick-fix to grounding problems on a ship is to drop a bare wire into the water with the other end attached to the recording equipment. However, a trial-and-error approach may be needed to resolve this.

Flow noise is a problem that causes artifacts in the recordings. Noise from water flow over the hydrophone and its mooring can create turbulence and small eddies (vortex shedding). These lead to fluctuating pressure around the hydrophone, which is sensed by the hydrophone and appears as noise in recordings. But this "noise" is not due to a traveling acoustic wave and hence not due to sound in the environment. It is an artifact. Flow noise is often a problem in rivers but also offshore (see flow noise marked in the spectrograms in Fig. 3.3 in Erbe et al. 2015). It can require the use of a shield or deflector, or placement of the hydrophone in a sheltered area.

Sound-recording acoustic tags are attached to marine animals to record their vocalizations and examine the effects of anthropogenic noise in the marine environment relative to animal generated sound. Flow noise (generated simply by water flowing around the tag) can be useful in this instance, as it can measure whale speed (von Benda-Beckmann et al. 2016; Fig. 3.9). However, interference by background noise is also a common problem. Unfortunately, survey vessels produce noise while operating. Therefore, to avoid unnecessary mechanical background noise during recordings, turn off any non-essential equipment (such as engines, pumps, filters, fans, generators, lights, refrigerators, winches, etc.). However, fishing, military, research, and whale-watching boat operators often are reluctant to do this. Alternatively, these vessel sounds can be filtered out during recording or analysis.

In rivers or shallow coastal areas, currents and tides transport sediment which may create noise. It may come as quite a shock when an entire recording is ruined by nonstop sand swishing



**Fig. 3.9** Non-animal generated noise can affect aquatic recordings adversely unless the research has a system in place that accounts for noise versus animal generated calls. Simply attaching a hydrophone or tag to a marine mammal can cause flow noise from water rushing around the attached object

back and forth over the hydrophone, creating noise between 10 Hz and 2 kHz (Erbe 2009). Perhaps more amusing shallow-water "mooring noise" occurred when a group of teenage girls swam over to the mooring, held on to the floats and sang ABBA songs for 20 minutes—very clearly recorded. The entire recording session had to be discarded (Erbe 2013).

Similarly, a hydrophone fixed to a ship, boat, buoy, or dock will bob up-and-down and produce spurious signals such as flow noise as the water passes the hydrophone and artifacts from hydrostatic pressure changes as the hydrophone changes its depth. The recording can be saturated with such signals. This noise can be reduced by suspending the hydrophone with a bungee cord, decoupling the floating hydrophone from the surface through a catenary line, or mounting the hydrophone on the seafloor (Fig. 3.10; also see Chap. 2, section on PAM systems). Another solution to reduce flow noise is to use a sonobuoy or an anti-heave buoy (see photograph in Chap. 4, section on sonobuoys). The long cable of the sonobuoy acts as a bungee cord to dampen vertical oscillations of the hydrophone. The sonobuoy is isolated from self-noise of the vessel, but will detect sounds from the vessel until it moves out of range.

Local sound propagation conditions will affect the recording (see Chap. 6 on sound propagation under water). It is important to measure and understand the sound speed profile in the study area to know the propagation pattern and range of a signal, which influence the recorded sound. For years, navies of the world measured sound speed profiles using disposable, battery-operated CTD (conductivity, temperature, depth) units, which were tossed into the ocean and data sent back to the ship as the unit fell in the water and unspooled a long copper wire. The units were not retrieved. Today, retrievable, digital CTD units are used. The sound speed profile may change over the course of a day—within the upper few meters below the sea surface. Turl and Thomas (1992) documented that a false killer whale (*Pseudorca crassidens*) echolocating during target-detection distance experiments in Kaneohe Bay, Hawaii, USA, consistently performed better during the morning than afternoon; i.e., the whale could detect the target at a greater distance during the morning. After taking CTD measurements prior to the morning and afternoon sessions, the researchers realized the water column, and thus sound speed profile, were very different between the two periods because or prevailing midday rains.

Sound propagation is particularly complicated in shallow water because of the close proximity of boundaries formed by the sea surface and seabed (Rogers and Cox 1988). Sound is reflected,

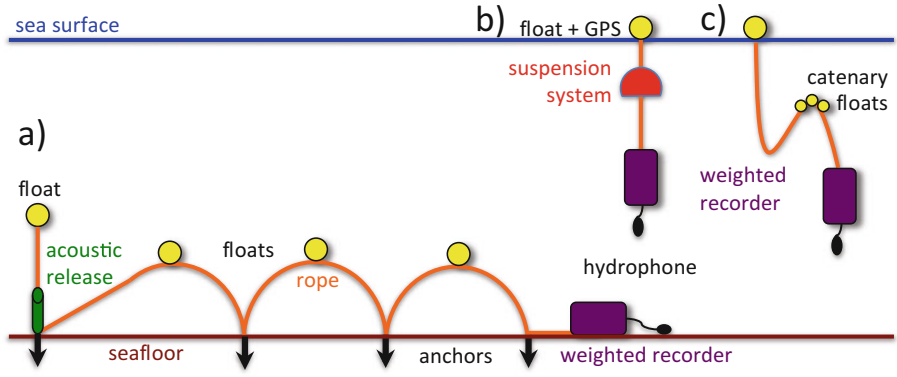



**Fig. 3.10** Mooring options to avoid noise artifacts: (**a**) recorder on the seafloor, (**b**) recorder suspended from a float via a bungee cord and drogue, and (**c**) recorder suspended via a catenary line (Erbe et al. 2019). © Erbe et al.; https://www.frontiersin.org/articles/10.3389/fmars.2019.00606/full. Published under a Creative Commons Attribution License (CC BY); https://creativecommons.org/licenses/by/4.0/

scattered, and absorbed at these boundaries. There is far more attenuation of low-frequency sounds in shallow water compared to deep water. Rogers and Cox (1988) suggested that the lowest frequency that could propagate in water less than 1 m deep was about 300 Hz, but this was strongly dependent on the nature of the seabed (sand, rock, or mud).

Ambient noise is an omnipresent issue and may mask the signals desired for recording (see Chap. 7 on soundscapes). Wind and precipitation create noise underwater from coastal to offshore regions. In polar regions, ice popping and cracking may dominate the soundscape. When a hydrophone was dropped in the ice-covered water next to a group of Antarctic Weddell seals (JAT, personal observations), music was heard from the radio-station at the New Zealand Research Base in Antarctica about 2 km away! Organisms from tiny snapping shrimp to enormous singing whales may also mask recordings of a target species. Ship noise is almost omnipresent in the world's oceans, so it can be difficult to obtain recordings of a target species in a quiet aquatic environment.

## 3.7 Considerations for Studies on Captive Animals

Because there are regulations on the housing and care of captive animals, research permit and IACUC requirements can be more detailed for research on captive species. However, often those regulations were written for laboratory animals used in medical research (mostly *Rattus* and *Mus*) and are not specified or applicable for wild animal research. For example, one of us (WLG) had to convince the university veterinarian to allow kangaroo rats (Heteromyidae, *Dipodomys*) to be housed using sandy desert soils instead of rat bedding so that these wild animals could properly sand-bathe and tunnel.

Zoos and aquaria support bioacoustical studies on a wide variety of species, including endangered species. Some benefits of studying captive animals in a zoo are that their history is usually known (i.e., wild caught vs. captive born, sex, age, reproductive history, relatedness to other animals, and health). Care should be taken to study healthy animals, as opposed to ill or rehabilitating animals, to best represent the acoustic abilities of their wild counterparts. However, burgeoning research by Therrien et al. (2012) indicated that changes in vocal behavior of bottlenose dolphins (*Tursiops truncatus*) and California sea lions (*Zalophus californianus*) actually could be used to indicate a health problem (Schwalm 2012). Moreover, captive animals, especially those that have been hand-reared or raised in a hatchery (such as salmon or sea bass) can show some degree of genetic selection, de-sensitization, and habituation to the presence of high levels of ambient sound. They can be much less responsive to sounds than wild animals.

Most zoos have noise created by loudspeaker announcements, music, shows, rides, or facility vehicles. Key events, such as hearing music for a show, or a vehicle delivering food, may affect animal behavior; therefore, studies should not be conducted during those times. Reminiscent of Ivan Pavlov in the 1890s experiment that dogs were being conditioned behaviorally (drooled) in response to being fed at the sound of a bell (conditioned response), researchers need to be aware of regular triggers to animal behavior. Of course, a common source of noise in captive studies is from visitors, keepers, and maintenance workers. If at all possible, it is best to conduct research before or after humans are near the study location (i.e., before or after the zoo is open). If possible, operation of air conditioners, furnaces, air-filters, and lights should be stopped, or minimized, to reduce or eliminate background sounds in recordings. Some facilities isolate their mechanical equipment in a separate building from the animals' environment; this greatly reduces noise exposure for the animals. A preliminary survey of noise in the animals' enclosure, using a sound pressure level meter, helps identify any particularly noisy or quiet areas.

Sometimes, ultrasonic noise or underwater noise can be present unbeknownst to zoo or aquarium staff. One of us (JAT, personal observations) provided two examples. In an underwater hearing study on a Pacific white-

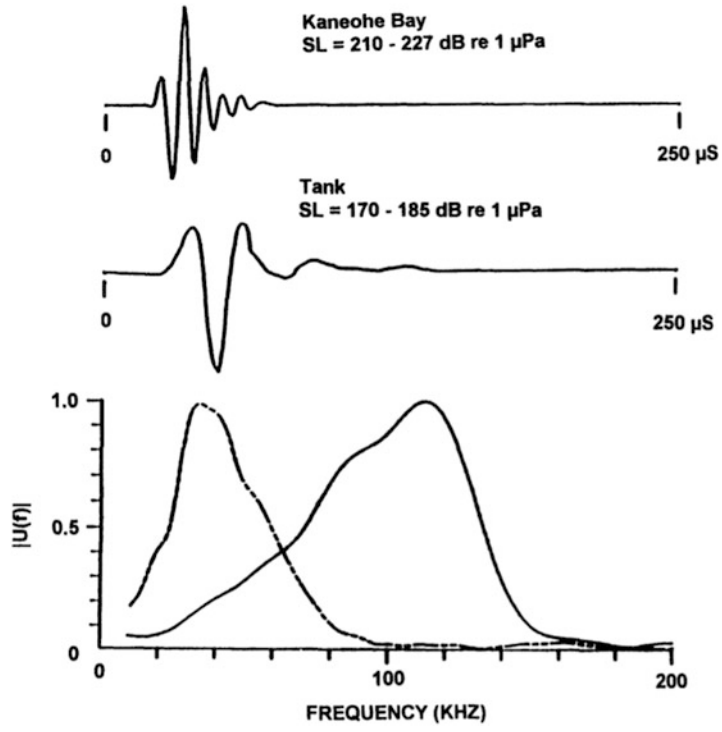



**Fig. 3.11** Waveforms and spectra of echolocation clicks of bottlenose dolphins in open ocean (Kaneohe Bay, Hawaii, USA) and in a tank. The spectrum of the click from the tank had a lower frequency peak at 40 kHz and a lower source level of 170–185 dB re 1 µPa m. Reprinted by permission from Springer Nature. Hearing by Whales and Dolphins, edited by W. W. L. Au, A. N. Popper, and R. R. Fay, pp. 364–408, Echolocation in dolphins, W. W. L. Au; https://doi.org/10.1007/978-1-4612-1150-1_9. © Springer Nature, 2000. All rights reserved

sided dolphin (*Lagenorhynchus obliquidens*) by Tremel et al. (1998), the test animal consistently reported hearing a 32-kHz signal at two different thresholds on different days. Spectrum analysis of the ambient noise in the pool revealed an intermittent noise near 32 kHz. So, on test days when the noise was present, the animal's threshold at this frequency was much lower than on test days when the noise was absent. Because the noise was ultrasonic, it was not known by staff or researchers. In another study by Therrien et al. (2012), 24-hour recordings of bottlenose dolphins detected an almost continuous banging noise in the water. Zoo staff were unaware of the noise and upon a diver's inspection of the pool, found a metal gate hinge that was broken and causing the banging sound. In both these examples, staff did not know about the noise, which could have been annoying to the animals and disturb bioacoustical research.

Researchers should understand the possible effects of the exhibit environment on the acoustic behavior of animals. For example, dolphins living in highly reverberant concrete pools echolocate less and at lower amplitudes than in the wild (Fig. 3.11) (Au 2000).

Today, exhibit designers incorporate irregular wall and floor surfaces in pools, indoor enclosures, and outdoor exhibits to minimize reverberations. Projecting a signal into a regularly shaped (e.g., round or square) pool with a flat bottom (e.g., during a hearing test) can set up standing waves, which result in a sound-field that dramatically changes with receiver location and frequency. A resonant pool amplifies sound at its resonance frequencies and dampens others, essentially distorting the signal desired by the researcher. While concrete walls in a zoo or aquarium are easy to construct and clean, they provide a reflective surface that often causes annoying, cave-like reverberations.

Particular issues are encountered when trying to perform hearing tests and sound exposure experiments with fish or invertebrates in water-filled tanks that are only a few meters in dimensions, or even smaller. The complexities

of the sound-field in small tanks were first pointed out by Parvulescu (1964) and recently discussed by Duncan et al. (2016), Grey et al. (2016), Rogers et al. (2016), and Popper and Hawkins (2018). Even in quite large tanks, the sound-field generated by even a simple sound source is transformed by interactions with boundaries (i.e., walls, floor of pool, and water surface) and can vary rapidly as a function of both space and frequency. The resulting sound-field can be difficult to model, or even characterize, and the sound-level can be very different from the natural environment. In particular, the levels of the particle motion components of the sounds (to which fish are sensitive) can be very high. Attempts at dampening reverberation by adding materials such as "horse hair" or bubble-wrap can be effective at high frequencies, but have little effect at the low frequencies to which fish are sensitive and where the sound wavelength often exceeds the dimensions of the tank (Popper and Hawkins 2018). In contrast, experiments performed in deep and open water allow the establishment of a relatively simple, well-controlled, and predictable sound-field (Hawkins 2014).

Grey et al. (2016) measured the sound-field in several large laboratory tanks and came to the following conclusions: 1) Tanks, even large ones, are not appropriate surrogates for open-water environments. 2) Tank wall-thickness is largely irrelevant. Walls backed by air essentially present a low impedance, and walls in contact with a solid foundation or ground present finite (non-rigid) impedance defined by the substrate materials. 3) Resonance of the tank walls can dominate underwater sound-field characteristics. 4) Lining the walls of a tank with acoustic absorbent material is futile, because the thicknesses required at low frequencies would leave no room for the fish. 5) Both the sound pressure and the particle motion of a sound need to be measured and checked for mutual validation by calculating the particle motion from pressure gradients. Special hydrophone systems, based on seismic accelerometers, are required to measure particle motion (see Chap. 2).

## 3.8 Digital File Format

Several file formats are available to save digital recordings. Digital file extensions include WAV, PCM, MP3, au, ram, MIDI, ogg, as well as others. It is best to record using uncompressed or WAV or PCM (Pulse Code Modulation) formats for faithful spectrum analysis.

MP3 is a digital audio-encoding format which uses data compression to reduce file size. It is a common audio-format for consumer audio and a de facto standard of digital audio-compression used for the transfer and playback of music. However, MP3 files and other compression methods are poor for spectrum analysis because compression only retains signals in a frequency band up to 16 kHz (i.e., the human hearing range). As a result, spectrum analysis using MP3 files is not trustworthy above 16 kHz. The psychoacoustic-based compression algorithms, in addition to limiting frequencies to below 16 kHz (and even less at higher compression ratios), discards fine details that cannot be heard by humans. Cuts introduced by compression appear as unpleasant "holes" in the spectrogram and can destroy details that could have meaning. However, MP3 files can be valuable for ecological monitoring of temporal and spatial patterns of well-known sounds.

A few digital recorders offer the Free Lossless Audio Codec (FLAC) format, which has less compression and reduces the storage space up to 50% without loss of detail. In addition, a few digital recorders employ a Direct Stream Digital (DSD) format; a proprietary system of digitally recreating audible signals for the Super Audio CD, using delta-sigma 1-bit A/D-converters at 2.8 or 5.6 MHz. Because of the intrinsic properties of the delta-sigma conversion made by the 1-bit A/D-converter, these recorders have the potential to record frequencies well beyond 100 kHz, but with increased noise at high frequencies. Spectrum analysis of recordings made in the DSD format is appropriate.

Waveform sound files (WAV; created by Microsoft) are perhaps the simplest of the common formats for storing audio samples. Unlike

MPEG and other compressed formats, WAV files and their derivatives (like the Broadcast Wave File, BWF) store samples "in the raw" where no pre-processing is used, other than formatting of data. When there is a choice of a recording file format, the WAV (or BWF) format should be selected, rather than the MP3 format.

With continuous recording, WAV files can become quite large and subsequently be difficult to handle with sound analysis software. For example, WAV recordings sampling at 96 kHz and 24 bit for 1 hour will occupy approximately 1 GB of storage capacity (96,000 samples/s × 24 bits × 1 byte/8 bits × 60 minutes × 60 s/minute = 1.04 GB). If monitoring is required for long periods, it is therefore important to select the appropriate sampling rate to conserve storage space. For example, if mid-frequency fish sounds are the main features of interest, then it can be appropriately sampled at only 22 kHz, or at an even lower sampling frequency. Several possible sampling frequencies and sometimes a choice of bit depth (16 or 24 bit) are available, but not on all recorders. Some recorders enable a limit to be placed on the maximum size of each recorded file. Alternatively, a recording protocol can be adopted to limit the length of each recording.

## 3.9 Data Storage

All storage media should be carefully labeled with who, what, where, and when. Each recording period should have a unique number. Creating a master catalog of recording numbers allows researchers to cross-reference metadata from a logbook.

Magnetic media, including magnetic tape (e.g., reel-to-reel, cassette, or DAT tapes), and computer hard drives require storage in a dry, dark area away from any type of magnetic field. Exposure to a magnet could erase data. If tapes are not played often, the tightly packed tape could "bleed through" from one segment to another, thus contaminating data. Therefore, converting old recordings on magnetic tape to modern storage is becoming urgent for data on historic soundscapes and animals not be lost.

When converting analog to digital formats, usually using an A/D-converter, the sampling frequency must be at least twice the highest frequency recorded and the recordist needs to make sure that the parameters of the storage medium are adequate for the task. There are a number of free software applications for conversion of analog to digital formats.

Storage of digital recordings can be done on hard drives, optical drives, solid-state memory, or an Internet cloud. Bluetooth (a wireless technology standard) provides reliable exchange of data between fixed and mobile devices over short distances. Bluetooth uses UHF radio waves that are effective at a short distance.

## 3.10 Archiving Recordings

Properly curated recordings are critically important for assessing changes in soundscapes, ambient noise, and animal presence/absence and acoustic behavior over time. For example, underwater recordings made by the US Navy off the coast of California indicated a steady increase in background noise levels in the ocean in the last 60 years (from the 1960s). Marie Poland Fish, an oceanographer and marine biologist, recorded and analyzed the sounds of more than 300 species of marine life, from mammals to mussels. Her work (described and spectrograms provided in Fish and Mowbray 1970) helped the US Navy to distinguish fish and other animal sounds from the sounds made by submarines and remains a primary source for analysis of marine fish sounds.

Recordings of humpback whale songs date back to the 1970s and continue to document annual changes in their song within different populations. Williams et al. (2013) studied the changing songs of male savannah sparrows (*Passerculus sandwichensis*) recorded over three decades (1980–2011) on Kent Island, New Brunswick, in the Bay of Fundy. Life-long recordings of songs of white-crowned sparrows (*Zonotrichia leucophrys)* found they memorize syllables they hear at 10–50 days of age and sing the same song throughout their life. In contrast, life-long recordings of northern

mockingbirds (*Mimus polyglottos)* found they add elements to their songs throughout their lives. Only long-term archival data could be used for analysis of these trends. In this time of global warming and accelerated ice melts, archived recordings from the polar regions might become instrumental in monitoring the rate of climate change (by quantifying ice-cracking noise) and the effects on soundscapes and ecology (Obrist et al. 2010). The take-home message here is that good research practices with solid documentation and data archiving allow for future knowledge generation.

## 3.11 Repositories of Bioacoustical Data

Hafner et al. (1997) noted that collections of animal recordings with ancillary data are rich sources of reference material for bioacoustical studies. Archiving analog data by converting to a digital format has played an essential role in preserving data for future use. Species-specific sounds from a variety of regions and times, with associated voucher specimens and metadata, are available for researchers at a number of organizations. All collections and their corresponding links were valid as of 13 June 2022.

In Europe, there is a long tradition of recording animal sounds, in particular bird songs, and many collections have been published on vinyl discs and CDs, mainly in France and the UK. In 1969, the British Library of Wildlife Sounds[2] established holdings of more than 160,000 well-documented field-recordings covering all classes of sound-producing animals from many regions. More than 10,000 species of invertebrates, insects, amphibians, reptiles, fishes, birds, and mammals, including many rare and threatened species. A large number of these recordings were made for radio by the BBC Natural History Unit. The British Library supported a citizen-science program to create a map of the UK coastal soundscape in 2015.[3] Other European online sound libraries include: Tierstimmen Archiv[4] (approximately 120,000 sound recordings; Museum für Naturkunde, Berlin, Germany) Xeno-Canto[5] (595,000 recordings from approximately 10,250 bird species Naturalis Biodiversity Center, Leiden, Netherlands), and FonoZoo[6] (11,657 recordings of 1621 animal species; Fonoteca Zoológica, Museo Nacional de Ciencias Naturales (CSIC), Madrid, Spain).

In the USA, the Macaulay Library[7] (Cornell Lab of Ornithology, Ithaca, NY, USA) archived older analog, digital, and video recordings. To date, their holdings are approximately 24 million photos, 915,000 audio and 192,000 video recordings available for researchers. The K. Lisa Yang Center for Conservation Bioacoustics[8] (Cornell Lab of Ornithology, Ithaca, NY, USA) is everything "bird" including citizen science and masterful guides and information in ornithology (including bird vocalization identification apps and bird cams). The Museum of Southwestern Biology[9] (University of New Mexico, Albuquerque, NM, USA) and Museum of Vertebrate Zoology[10] (University of California, Berkeley, CA, USA) have hundreds of thousands of cataloged natural history journals and voucher specimens and began to associate avian vocalizations with voucher specimens in the 2000s. These museum collections have shown a desire to include bat call libraries before 2023. The Watkins Sound Library[11] (Woods Hole Oceanographic Institution, Woods Hole, MA, USA) provides particularly good collections of marine mammal sounds with a highlighted "Best of" cuts section that contains 1694 sound

[2] https://www.bl.uk/collection-guides/wildlife-and-environmental-sounds; accessed 13 June 2022

[3] https://www.bl.uk/sounds-of-our-shores

[4] http://www.tierstimmenarchiv.de/

[5] https://www.xeno-canto.org/

[6] http://www.fonozoo.com/index_eng.php

[7] http://macaulaylibrary.org

[8] https://www.birds.cornell.edu/ccb/

[9] https://arctosdb.org/; http://www.msb.unm.edu/

[10] http://mvz.berkeley.edu/General_Information.html

[11] https://cis.whoi.edu/science/B/whalesounds/index.cfm




**Fig. 3.12** Commercial companies and others market sounds of animals and soundscapes recorded by researchers such as Bernie Krause. Recording and analyzing natural sound is fulfilling and insightful, and can be a profound source for generating knowledge. Left photo by the authors; right photo, "Capturing the sounds of the lake" by S. Shiller; https://www.flickr.com/photos/12289718@N00/9454414945; licensed under CC BY 2.0; https://creativecommons.org/licenses/by/2.0/

cuts deemed to be of higher sound quality and lower noise from 32 different marine mammal species.

Several commercial companies market LPs and CDs of nature sounds. Bernie Krause[12] (Wild Sanctuary, Glen Ellen, CA, USA; Fig. 3.12) is unique among researchers, commercial ventures, and artists. From the Wild Sanctuary website, "The Wild Sanctuary Audio Archive represents a vast and important collection of whole-habitat field recordings and precise metadata dating from the late 1960s. This unique bioacoustic resource contains marine and terrestrial soundscapes representing the voices of living organisms from larvae to large mammals and the numerous tropical, temperate and Arctic biomes from which they come. The catalog currently contains over 4500 hours of wild soundscapes and in excess of 15,000 identified life forms." The acoustic world is not only at our finger tips, but the world is becoming available for all to hear.

[12] http://www.wildsanctuary.com/

## 3.12 Summary

As with other areas of science, good practices for bioacoustical research, as well as an awareness of the ethical implications of that research, should be employed. This chapter provides a list of considerations for terrestrial, aquatic, and captive studies—a list that will doubtlessly be improved as technology and access to the acoustic world improves. No longer is large, heavy, and expensive equipment necessary to make high-quality, meaningful acoustic recordings. Acoustic data are important beyond the immediate scope of a project, but data must be well documented with metadata (including field notes and ancillary information) and stored in a way that they are preserved and accessible for future research. The importance of a well-designed data sheet for easy data entry and analysis is also discussed along with special considerations for study design. Playbacks of sounds to animals are commonly used by bioacousticians and procedures for playbacks and controls are recommended.

Several sound libraries are publicly available for research. These facilities have invested a great

deal of time in transferring analog recordings to digital formats for more permanent preservation. CDs of animal and nature sounds are now commercially available. Archives are useful for education and research. As we evaluate current hypotheses related to global warming, perhaps we can hear the world change.

## 3.13 Additional Resources

- Sound recording tips from eBird: https://www.macaulaylibrary.org/how-to/recording-techniques/
- Bioacoustics equipment and field techniques, Centro Interdisciplinare di Bioacustica e Ricerche Ambientali, Università degli Studi di Pavia: http://www.unipv.it/cibra/edu_equipment_uk.html
- Manual on Field Recording Techniques and Protocols for All Taxa Biodiversity Inventories and Monitoring (Eymann et al. 2010): https://issuu.com/ysamyn/docs/abctaxa_vol_8_part1_lr

All web resources were last accessed 13 June 2022.

## References


Amorim MCP, Simões JM, Fonseca PJ, Turner GF (2008) Species differences in courtship acoustic signals among five Lake Malawi cichlid species (*Pseudotropheus* spp.). Fish Biol 72:1355–1368. https://doi.org/10.1111/j.1095-8649.2008.01802.x

Au WWL (2000) Chapter 9: Echolocation in dolphins. In: WWL A, Popper AN, Fay RR (eds) Hearing by whales and dolphins. Springer-Verlag, New York, pp 364–408

Beecher MD, Burt JM (2004) The Role of Social Interaction in Bird Song Learning. Curr Dir Psychol Sci 13: 224–228. https://doi.org/10.1111/j.0963-7214.2004.00313.x

Borkin KM, Smith DHV, Shaw WB, McQueen JC (2019) More traffic, less bat activity: the relationship between overnight traffic volumes and *Chalinolobus tuberculatus* activity along New Zealand highways. Acta Chiropterologica 21:321–329. https://doi.org/10.3161/15081109ACC2019.21.2.007

Carles F, Torre I, Arrizabalaga A (2007) Comparison of Sampling Methods for Inventory of Bat Communities. J Mammal 88:526–533. https://doi.org/10.1644/06-MAMM-A-135R1.1

Charrier I, Mathevon H, Aubin T (2013) Bearded seal males perceive geographic variation in their trills. Behav Ecol Sociobiol 67(10):1679–1689. https://doi.org/10.1007/s00265-013-1578-6

Clark CW, Clark JM (1980) Sound playback experiments with southern right whales (*Eubalaena australis*). Science 20:663–665

Davis J (1958) Singing behavior and the gonad cycle of the Rufous-sided towhee. Condor 60:308–336

Deecke V (2003) Seals are guided by voices. Discover April p. 17.

Duncan AJ, Lucke K, Erbe C, McCauley RD (2016) Issues associated with sound exposure experiments in tanks. Proc Meet Acoust 27(1):070008. https://doi.org/10.1121/2.0000280

Emlen ST (1972) An Experimental Analysis of the Parameters of Bird Song Eliciting Species Recognition. Anim Behav 41:130–171

Erbe C (2009) Underwater noise from pile driving in Moreton Bay, Qld. Acoust Aust 37(3):87–92

Erbe C (2013) Underwater noise of small personal watercraft (jet skis). J Acoust Soc Am 133(4):EL326–EL330. https://doi.org/10.1121/1.4795220

Erbe C, Verma A, McCauley R, Gavrilov A, Parnum I (2015) The marine soundscape of the Perth Canyon. Prog Oceanogr 137:38–51. https://doi.org/10.1016/j.pocean.2015.05.015

Erbe C, Marley S, Schoeman R, Smith JN, Trigg L, Embling CB (2019) The effects of ship noise on marine mammals--A review. Front Mar Sci 11. https://doi.org/10.3389/fmars.2019.00606

Eymann J, Degreef J, Hauser C, Monje JC, Samyn Y, VanDerSpiegel D (2010) Manual on field recording techniques and protocols for all taxa biodiversity inventories and monitoring. Abc Taxa series, The Belgian Development Corporation, Brussels, Belgium http://www.abctaxa.be ISSN 1784-1291.

Filatova OA, Fedutin ID, Burdin AM (2011) Responses of Kamchatkan fish-eating killer whales to playbacks of conspecific calls. Mar Mam Sci 27(2):E26–E42

Fish MP, Mowbray WH (1970) Sounds of Western North Atlantic fishes. A reference file of biological underwater sounds. The John Hopkins Press, Baltimore, MA, USA, 207 p

Gannon WL (1999) *Tamias siskiyou*. In: Wilson DE, Ruff S (eds) Complete book of North American Mammals. Smithsonian Institution Press, Washington, DC

Gannon WL (2014) Integrating research ethics with graduate education in Geography. J Geography High Ed 38:481–499. https://doi.org/10.1080/03098265.2014.958656

Gluck JP (1998) The death of a vagrant bird. In: Orlans BF, Beauchamp TL, Dresser R, Morton D, Gluck JP (eds) The human use of animals; Case studies in ethical choice. Oxford University Press, New York, pp 191–208

Grey MD, Rogers PH, Popper AN, Hawkins AD, Fay RR (2016) Large tank acoustics: How big is big enough? In: Popper AN, Hawkins A (eds) The effects of noise on aquatic life II, Advances in experimental Medicine and biology. Springer Science + Business

Media, New York. https://doi.org/10.1007/978-1-4939-2981-8_43
Hafner MS, Gannon WL, Salazar-Bravo J, Alvarez-Castaneda ST (1997) Mammal collections of the Western Hemisphere: a survey and directory of existing collections. In: American Society of Mammalogists. Allen Press, Lawrence KS, p 97. ISBN 0-89338-055-5. http://www.mammalsociety.org/uploads/committee_files/collsurvey.pdf
Hawkins AD (2014) Examining fish in the sea: a European perspective on fish hearing experiments. In: Popper AN, Fay RR (eds) Perspectives on auditory research. Springer 247 Handbook of Auditory Research, p 50. https://doi.org/10.1007/978-1-4614-9102-6_14
Hawkins AD, Roberts L, Cheesman S (2014) Responses of free-living coastal pelagic fish to impulsive sounds. J Acoust Soc Am 135(5):3101–3116
Hayes SA, Kumar A, Costa DP, Mellinger DK, Harvey J (2004) Evaluating the function of the male harbour seal, *Phoca vitulina*, roar through playback experiments. Anim Behav 67:1133–1139. https://doi.org/10.1016/j.anbehav.2003.06.019
Kissner KJ, Forbes MR, Secoy DM (1997) Rattling behavior of prairie rattlesnakes (*Crotalus viridis viridis*, Viperidae) in relation to sex, reproductive Status, Body Size, and Body Temperature. Ethology 103(12):1042–1050. https://doi.org/10.1111/j.1439-0310.1997.tb00146.x
Kunz TH, Parsons S (eds) (2009) Ecological and behavioral methods for the study of bats, 2nd edn. The John Hopkins Press, Baltimore, MD, p 920
Lindseth AV, Lobel PS (2018) Underwater soundscape monitoring and fish bioacoustics: a review. Fishes 3:1–15. https://doi.org/10.3390/fishes3030036
Mobley JR Jr, Herman IM, Frankel AS (1988) Responses of wintering humpback whales (*Megaptera novaeangliae*) to playback of recordings of winter and summer vocalisations and of synthetic sound. Behav Ecol Sociobiol 23:211–223
Morton SL, Morton ES (1998) Sound playback studies. In: Evans CS (ed) Animal acoustic communication. Springer-Verlag, Berlin, pp 323–352
Myrberg AA Jr, Riggio RJ (1985) Acoustically mediated individual recognition by a coral reef fish (*Pomacentrus partitus*). Anim Behav 33:411–416
Obrist MK, Pavan G, Sueur J, Riede K, Llusia D, Márquez R (2010) Bioacoustic approaches in biodiversity inventories. In: Manual on field recording techniques and protocols for all taxa biodiversity inventories, Abc Taxa series, vol 8, http://www.abctaxa.be, ISSN 1784-1291 edn. The Belgian Development Corporation, Brussels, pp 68–99
Parvulescu A (1964) Problems of propagation and processing. In: Tavolga WN (ed) Marine bio-acoustics. Pergamon Press, Oxford, pp 87–100
Popper AN, Hawkins AD (2018) The importance of particle motion to fishes and invertebrates. J Acoust Soc Am 143(1):470–488. https://doi.org/10.1121/1.5021594
Prat Y, Taub M, Yovel Y (2016) Everyday bat vocalizations contain information about emitter. Nature 39419(2016):1. https://doi.org/10.1038/srep39419
Rogers PH, Cox M (1988) Underwater sound as a biological stimulus. In: Atema J, Fay RR, Popper AN, Tavolga WN (eds) Sensory biology of aquatic animals. Springer-Verlag, New York, pp 131–149
Rogers PH, Hawkins AD, Popper AN, Fay RR, Grey MD (2016) Parvulescu revisited: small tank acoustics for bioacousticians. In: Popper AN, Hawkins A (eds) The effects of noise on aquatic life II, Advances in experimental medicine and biology. Springer Science + Business Media, New York. https://doi.org/10.1007/978-1-4939-2981-8_43
Schwalm A (2012) Analysis of aerial vocalizations of California sea lions in rehabilitation as an indicator of health. PhD Dissertation, Western Illinois University.
Sherwin RE, Gannon WL, Haymond S (2000) The efficacy of acoustic techniques to infer differential use of habitat by bats. Acta Chiropterologica 2(2):145–153
Simpson SD, Radford AN, Tickle EJ, Meekan MG, Jeffs AG (2011) Adaptive avoidance of reef noise. PLoS One 6(2):e16625. https://doi.org/10.1371/journal.pone.0016625
Slobodchikoff CN, Ackers SH, Van Ert M (1998) Geographical variation in alarm calls of Gunnison's prairie dogs. J Mammal 79:1265–1272
Slobodchikoff CN, Perla BS, Verdolin JL (2009) Prairie dogs: Communication and community in an animal society. Harvard University Press, Boston, MA
Smith AT, Nagy JD, Millar CI (2016) Behavioral ecology of American pikas (*Ochotona princeps*) at Mono Craters, California: Living on the edge. West N Am Nat 76(4):459–484
Therrien SC, Thomas JA, Therrien RE, Stacey R (2012) Time of day and social change affects underwater sound production of bottlenose dolphins (*Tursiops truncatus*) at the Brookfield Zoo. Aquat Mamm 38(1):65–75
Thomas JA, DeMaster DP (1982) An acoustic technique for determining haulout pattern in leopard (*Hydrurga leptonyx*) and crabeater (*Lobodon carcinophagus*) seals. Can J Zool 60(8):2028–2031
Thomas JA, Zinnel KC, Ferm LM (1983) Analysis of Weddell seal (*Leptonychotes weddelli*) vocalizations using underwater playbacks. Can J Zool 61:1448–1456
Thomas JA, Friel B, Yegge S (2016) Restoring duetting behavior in a mated pair of buffy cheeked gibbons after exposure to construction noise at a zoo through playbacks of their own sounds ASA abstract Dec 2016.
Tremel DP, Thomas JA, Ramirez KT, Dye GS, Bachman WA, Orban AN, Grimm KK (1998) Underwater hearing sensitivity of a Pacific white-sided dolphin, *Lagenorhynchus obliquidens*. Aquat Mamm 24(2):63–69
Turl CW, Thomas JA (1992) Possible relationship between oceanographic conditions and long-range target detection by a false killer whale. In: Thomas JA,

Kastelein RA, Supin AY (eds) Marine Mammal Sensory Systems. Plenum Press, New York, pp 421–432. 773 pp. ISBN 9780306443510

Tyack P (1983) Differential responses of humpback whales, *Megaptera novaeangliae,* to playback of song or social sounds. Behav Ecol Sociobiol 13:49–55

von Benda-Beckmann AM, Wensveen PJ, Samarra FIP, Beerens SP, Miller PJO (2016) Separating underwater ambient noise from flow noise recorded on stereo acoustic tags attached to marine mammals. J Exp Biol 219, 2271–2275. https://doi.org/10.1242/jeb.133116.

White M (2013) The ethical flap over birdsong apps. National Geographic, https://www.nationalgeographic.com/news/2013/6/130614-bird-watching-birdsong-smartphone-app-ethics/

Williams H, Levin II, Ryan-Norris D, Newman AEM, Wheelwright NT (2013) Three decades of cultural evolution in Savannah Sparrow songs. Anim Behav 85(1):213. https://doi.org/10.1016/j.anbehav.2012.10.028

Yegge SA (2012) Playbacks of conspecific vocalizations and music to Gabriella's crested gibbons at Niabi Zoo to restore their duetting behavior Western Illinois University. PhD Dissertation.

Young S, Carr A, Jones G (2018) CCTV enables the discovery of new barbastelle (*Barbastella barbastellus*) vocalizations and activity pattern near a roost. Acta Chiropterologica 20:262–272

Zelick R, Mann DA, Popper AN (1999) Acoustic communication in fishes and frogs. In: Fay RR, Popper AN (eds) Comparative hearing: fish and amphibians. Springer-Verlag, New York, pp 363–411

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



# Introduction to Acoustic Terminology and Signal Processing

4

Christine Erbe, Alec Duncan, Lauren Hawkins, John M. Terhune, and Jeanette A. Thomas

## 4.1 What Is Sound?

Most people think of sound as something they can hear, such as speech, music, bird song, or noise from an overflying airplane. There has to be a source of sound, such as another person, an animal, or a train. The sound then travels from the source through the air to our ears. Acoustics is the science of sound and includes the generation, propagation, reception, and effects of sound. The more scientific definition of sound refers to an oscillation in pressure and particle displacement that propagates through an acoustic medium (American National Standards Institute 2013; International Organization for Standardization 2017). Sound can also be defined as an auditory sensation that is evoked by such oscillation (American National Standards Institute 2013), however, more general definitions do not require a human listener, do allow for an animal receiver, or don't require a receiver at all.

Not all sounds produce an auditory sensation in humans. For example, ultrasound refers to sound at frequencies above 20 kHz, while infrasound refers to frequencies below 20 Hz. These definitions are based on the human hearing range of 20 Hz – 20 kHz (American National Standards Institute 2013). While sound outside of the human hearing range is inaudible to humans, it may be audible to certain animals. For example, dolphins hear well into high ultrasonic frequencies above 100 kHz. Also, inaudible doesn't mean that the sound cannot cause an effect. For example, infrasound from wind turbines has been linked to nausea and other symptoms in humans (Tonin 2018). As well, the effects of ultrasound on humans have been of concern (Parrack 1966; Acton 1974; Leighton 2018).

Noise is also sound, but typically considered unwanted. It therefore requires a listener and includes an aspect of perception. Whether a sound is perceived as noise depends on the listener, the situation, as well as acquired cognitive and emotional experiences with that sound. Different listeners might perceive sound differently and classify different sound as noise. One person's music is another person's noise. Noise could be the sound near an airport that has the potential to mask speech. It could be the ambient noise at a recording site and encompass sound from a multitude of sources near and far. It could be the recorder's electric self-noise (see also American National Standards

Jeanette A. Thomas (deceased) contributed to this chapter while at the Department of Biological Sciences, Western Illinois University-Quad Cities, Moline, IL, USA

C. Erbe (✉) · A. Duncan · L. Hawkins
Centre for Marine Science and Technology, Curtin University, Perth, WA, Australia
e-mail: c.erbe@curtin.edu.au; A.J.Duncan@curtin.edu.au

J. M. Terhune
Department of Biological Sciences, University of New Brunswick, Saint John, NB, Canada
e-mail: terhune@unb.ca

© The Author(s) 2022

C. Erbe, J. A. Thomas (eds.), *Exploring Animal Behavior Through Sound: Volume 1*,
https://doi.org/10.1007/978-3-030-97540-1_4

Institute 2013; International Organization for Standardization 2017). In contrast to noise, a signal is wanted, because it conveys information.

There are many ways to describe, quantify, and classify sounds. One way is to label sounds according to the medium in which they have traveled: air-borne, water-borne, or structure-borne (also called substrate-borne or ground-borne). For example, scientists studying bat echolocation work with air-borne sound. Those looking at the effects of marine seismic survey noise on baleen whales work with water-borne sounds. Some of the sound may have traveled as a structural vibration through the ground and is therefore referred to as structure-borne. Just as earthquakes can be felt on land, submarine earthquakes can be sensed by benthic organisms on the seafloor. In both cases, the sound is structure-borne (Dziak et al. 2004). Sound can cross from one medium into another. The sound of airplanes is generated and heard in air but also transmits into water where it may be detected by aquatic fauna (e.g., Erbe et al. 2017b; Kuehne et al. 2020).

Another way of grouping sounds is by their sources: geophysical, biological, or anthropogenic. Geophysical sources of sound are wind, rain, hail, breaking waves, polar ice, earthquakes, and volcanoes. Biological sounds are made by animals on land, such as insects, birds, and bats, or by animals in water, such as invertebrates, fishes, and whales. Anthropogenic sounds are made by humans and stem from airplanes, cars, trains, ships, and construction sites. The distinction by source type is common in the study of soundscapes. These comprise a geophony, biophony, and anthropophony.

The following sections explain some of the physical measurements by which sounds can be characterized and quantified. The terminology is based on international standards (including, International Organization for Standardization 2007, 2017; American National Standards Institute 2013).

## 4.2 Terms and Definitions

### 4.2.1 Units

A wide (and confusing) collection of units can be found in early books and papers on acoustics, but the units now used for all scientific work are based on the International System of Units, better known as the SI system (Taylor and Thompson 2008). In this system, a unit is specified by a standard symbol representing the unit itself, and a multiplier prefix representing a power of 10 multiples of that unit. For example, the symbol μPa (pronounced micro pascal) is made up of the multiplier prefix μ (micro), representing a factor of $10^{-6}$ (one one-millionth) and the symbol Pa (pascal), which is the SI unit of pressure. So, a measured pressure given as 1.4 μPa corresponds to 1.4 times $10^{-6}$ Pa or 0.0000014 Pa. The SI base units are listed in Table 4.1. Other quantities and their units result from quantity equations that are

**Table 4.1** SI base units (length, mass, time, electric current, temperature, luminous intensity, and amount of substance) and example derived units (frequency, pressure, energy, and power)

| Quantity | Unit name | Unit symbol | Expressed in terms of base units |
|---|---|---|---|
| Length | meter | m | |
| Mass | kilogram | kg | |
| Time | second | s | |
| Electric current | ampere | A | |
| Temperature | kelvin | K | |
| Luminous intensity | candela | cd | |
| Amount of substance | mole | mol | |
| Frequency | hertz | Hz | 1 / s |
| Pressure | pascal | Pa | kg / (m $s^2$) |
| Energy | joule | J | kg $m^2$ / $s^2$ |
| Power | watt | W | kg $m^2$ / $s^3$ |

**Table 4.2** SI multiplier prefixes

| Prefix | Symbol | Factor | Prefix | Symbol | Factor |
|---|---|---|---|---|---|
| deci | d | $10^{-1}$ | deka | da | $10^{1}$ |
| centi | c | $10^{-2}$ | hecto | h | $10^{2}$ |
| milli | m | $10^{-3}$ | kilo | k | $10^{3}$ |
| micro | μ | $10^{-6}$ | mega | M | $10^{6}$ |
| nano | n | $10^{-9}$ | giga | G | $10^{9}$ |
| pico | p | $10^{-12}$ | tera | T | $10^{12}$ |

based on these base quantities. The SI multiplier prefixes that go along with these units are listed in Table 4.2. Note that unit names are always written in lowercase. However, if the unit is named after a person, then the symbol is capitalized, otherwise the symbol is also lowercase. Examples for units named in honor of a person are kelvin [K], pascal [Pa], and hertz [Hz].

### 4.2.2 Sound

Sound refers to a mechanical wave that creates a local disturbance in pressure, stress, particle displacement, and other quantities, and that propagates through a compressible medium by oscillation of its particles. These particles are acted upon by internal elastic forces. Air and water are both fluid acoustic media and sound in these media travels as longitudinal waves (also called pressure or P-waves). A common misconception is that the air or water particles travel with the sound wave from the source to a receiver. This is not the case. Instead, individual particles oscillate back and forth about their equilibrium position. These oscillations are coupled across individual particles, which creates alternating regions of compressions and rarefactions and which allows the sound wave to propagate (Fig. 4.1[1]). The line along which the particles oscillate is parallel (or longitudinal) to the direction of propagation of the sound wave in the case of longitudinal waves.

Rock is a solid medium and here, vibration travels as both longitudinal (also called pressure or P-waves) and transverse waves (also called shear or S-waves). In S-waves, the particles oscillate perpendicular to the direction of propagation. It is again because of the coupling of particles, that the wave propagates. P-waves travel faster than S-waves so that P-waves arrive before S-waves. The P therefore also stands for "primary" and S for "secondary."

### 4.2.3 Frequency

Frequency refers to the rate of oscillation. Specifically, it is the rate of change of the phase of a sine wave over time, divided by $2\pi$. Here, phase refers to the argument of a sine (or cosine) function. It denotes a particular point in the cycle of a waveform. Phase changes with time. Phase is measured as an angle in radians or degrees. Phase is a very important factor in the interaction of one wave with another. Phase is not normally an audible characteristic of a sound wave, though it can be in the case of very-low-frequency sounds.

A simpler concept of frequency of a sine wave, as shown in Fig. 4.1, is the number of cycles per second. A full cycle lasts from one positive peak to the next positive peak. To determine the frequency, count how many full cycles and fractions thereof occur in 1 s. Note that pitch is an attribute of auditory sensation and while it is related to frequency, it is used in human auditory perception as a means to order sounds on a musical scale. As

[1] Dan Russell's animations of particle motion during acoustic wave propagation: https://www.acs.psu.edu/drussell/Demos/waves-intro/waves-intro.html, of the amplitude at a fixed location: https://www.acs.psu.edu/drussell/Demos/wave-x-t/wave-x-t.html, and of longitudinal and transverse waves: https://www.acs.psu.edu/drussell/Demos/waves/wavemotion.html; accessed 12 October 2020.

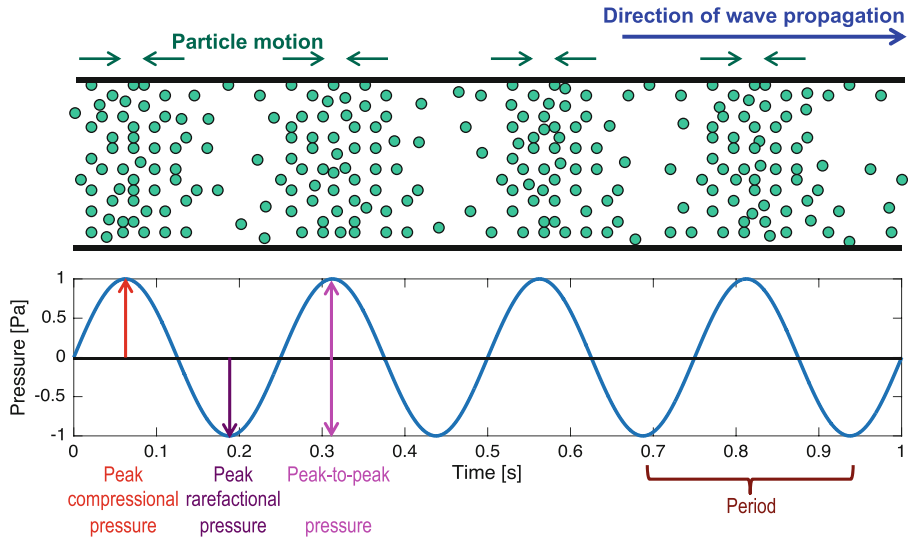



**Fig. 4.1** A sinusoidal sound wave having a peak pressure of 1 Pa, a peak-to-peak pressure of 2 Pa, a root-mean-square pressure of 0.7 Pa, a period of 0.25 s, and a frequency of 4 Hz. The top plot indicates the motion of the particles of the medium; they undergo coupled oscillations back and forth, so that the sound wave propagates to the right. At regions of compression, the pressure is high; at regions of rarefaction, it is low. The bottom plot shows the change in pressure over time at a fixed location. While the plots are lined up, the horizontal axes of the top and bottom plots are space and time, respectively

we know very little about auditory perception in animals, the term pitch is not normally used in animal bioacoustics.

The symbol for frequency is $f$ and the unit is hertz [Hz] in honor of Heinrich Rudolf Hertz, a German physicist who proved the existence of electromagnetic waves. Expressed in SI units, 1 Hz = 1/s.

The fundamental frequency (symbol: $f_0$; unit: Hz) of an oscillation is the reciprocal of the period. The period (symbol: $\tau$; unit: s) is the duration of one cycle and is related to the fundamental frequency as (see Fig. 4.1):

$$\tau = \frac{1}{f_0}$$

The wavelength (symbol: λ; unit: m) of a sine wave measures the spatial distance between two successive "peaks" or other identifiable points on the wave.

A sound that consists of only one frequency is commonly called a pure tone. Very often, sounds contain not only the fundamental frequency but also harmonically related overtones. The frequencies of overtones are integer multiples of the fundamental: 2 $f_0$, 3 $f_0$, 4 $f_0$, ... Beware that there are two schemes for naming these tones: $f_0$ can be called either the fundamental or the first harmonic. In the former case, 2 $f_0$ becomes the first overtone, 3 $f_0$ the second overtone, etc. In the latter case, 2 $f_0$ becomes the second harmonic, 3 $f_0$ the third harmonic, etc.

Musical instruments produce harmonics, which determine the characteristic timbre of the sounds they produce. For example, it is the differences in harmonics that make a flute sound unmistakably different from a clarinet, even when they are playing the same note. Animal sounds also often have harmonics as they use similar basic mechanisms to musical instruments. Most mammals have string-like vocal cords and birds have string-like syrinxes. Fish have muscles that contract around a swim bladder to produce percussive-type sounds. Insects and invertebrates stridulate or rub body parts together to produce a percussive sound.

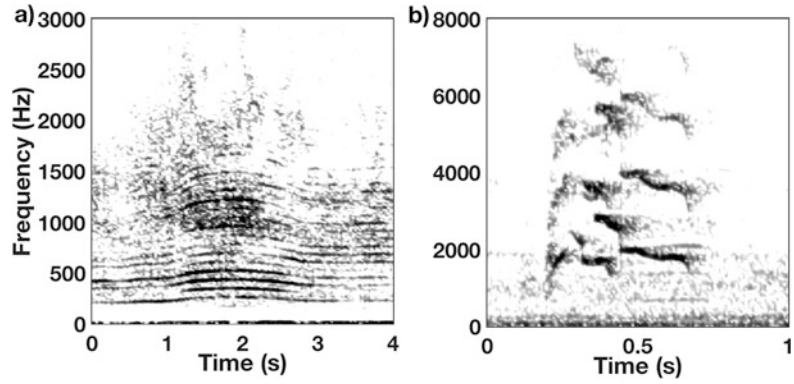



**Fig. 4.2** Spectrograms of (**a**) a jet ski recorded under water Erbe 2013 and (**b**) a Carnaby's Cockatoo (*Calyptorhynchus latirostris*) whistle, both displaying frequency modulation

The frequency or frequencies of a sound may change over time, so that frequency is a function of time: $f(t)$. This is called frequency modulation (abbreviation: FM). If the frequency increases over time, the sound is called an upsweep. If the frequency decreases over time, the sound is called a downsweep. Sounds without frequency modulation are called continuous wave. The sound of jet skis under water is frequency-modulated due to frequent speed changes (Erbe 2013). Whistles of animals such as birds or dolphins (e.g., Ward et al. 2016) are commonly frequency-modulated and often exhibit overtones (Fig. 4.2).

The acoustic features of frequency-modulated sounds such as whistles can identify the species, population, and sometimes individual animal that made them (e.g., Caldwell and Caldwell 1965). Such characteristic features include the start frequency, end frequency, minimum frequency, maximum frequency, duration, number of local extrema, number of inflection points, and number of steps (e.g., Marley et al. 2017). The start frequency is the frequency at the beginning of the fundamental contour, the end frequency is the frequency at the end of the fundamental contour (Fig. 4.3). The minimum frequency is the lowest frequency of the fundamental contour and the maximum frequency is the highest. Duration measures how long the whistle lasts. Extrema are points of local minima or maxima in the contour. At a local minimum, the contour changes from downsweep to upsweep; at a local maximum, it changes from upsweep to downsweep. Mathematically, the first derivative of the whistle contour with respect to time is zero at a local extremum, and the second derivate is a positive number in the case of a minimum or a negative number in the case of a maximum. At an inflection point, the curvature of the contour changes from clockwise to counter-clockwise or vice versa. Mathematically, the first derivative of the whistle contour with respect to time exhibits a local extremum and the second derivative is zero at an inflection point. Steps in the contour are discontinuities in frequency. There is no temporal gap but the contour jumps in frequency. The frequency measurements are taken from the fundamental contour. The duration, number of local extrema, number of inflection points, and number of steps are the same in fundamental and overtones and can therefore be measured from any harmonic contour. This is beneficial if the fundamental is partly masked by noise.

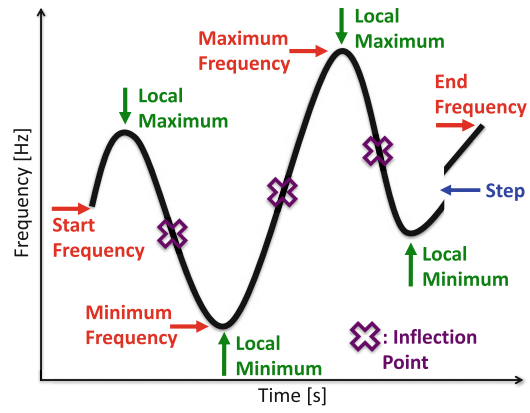



**Fig. 4.3** Spectrogram of a frequency-modulated sound, identifying characteristic features

### 4.2.4 Pressure

Atmospheric pressure is the static pressure at a specified height above ground and is due to the weight of the atmosphere above. Similarly, hydrostatic pressure is the static pressure at a specified depth below the sea surface and is due to the weight of the water above plus the weight of the atmosphere.

Sound pressure (or acoustic pressure) is caused by a sound wave. Sound pressure (symbol: $p$; unit: Pa) is dynamic pressure; it varies with time $t$ (i.e., $p$ is a function of $t$: $p(t)$). It is a deviation from the static pressure and defined as the difference between the instantaneous pressure and the static pressure. Air-borne sound pressure is measured with a microphone, water-borne sound pressure with a hydrophone. The unit of pressure is pascal [Pa] in honor of Blaise Pascal, a French mathematician and physicist. Some of the superseded units of pressure are bar and dynes per square centimeter, which can be converted to pascal: 1 bar $= 10^6$ dyn/cm$^2$ $= 10^5$ Pa. Mathematically, pressure is defined as force per area. Pascal in SI units is

$$1\ \text{Pa} = 1\ \text{N/m}^2 = 1\ \text{J/m}^3 = 1\ \text{kg/}\left(\text{m s}^2\right)$$

where N symbolizes newton, the unit of force, and J symbolizes joule, the unit of energy.

The pressure in Fig. 4.1 follows a sine wave: $p(t) = A \sin(2\pi ft)$, where $A$ is the amplitude and $f$ the frequency. In the example of Fig. 4.1, $A = 1$ Pa, $f = 4$ Hz. In general terms, the amplitude is the magnitude of the largest departure of a periodically varying quantity (such as sound pressure or particle velocity, see Sect. 4.2.8) from its equilibrium value. The magnitude is always positive and commonly symbolized by two vertical bars: $|p(t)|$. These are the same values as $p(t)$, but without the sign (i.e., the magnitude is always positive). The amplitude may not always be a constant. When it changes as a function of time $A(t)$, the signal undergoes amplitude modulation (abbreviation: AM).

The signal in Fig. 4.4 is both amplitude- and frequency-modulated:

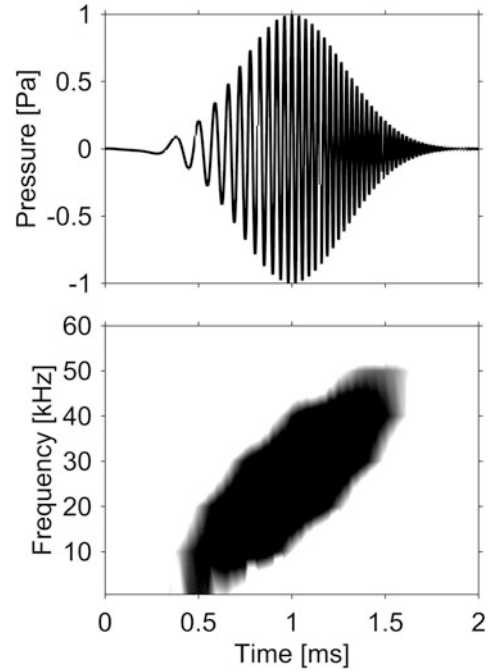



**Fig. 4.4** Gabor click similar to a beaked whale click. The signal is based on a sine wave; the amplitude is modulated by a Gaussian function, and the frequency is swept up with time. The corresponding spectrogram is shown in the bottom panel

$$p(t) = A(t)\ \sin\left(2\,\pi f(t) \times t\right)$$

The amplitude function changes exponentially with time:

$A(t) = e^{-(t-t_0)^2/2\sigma^2}$, where the peak occurs at $t_0 = 1$ ms, and $\sigma$ is the standard deviation of the Gaussian envelope. Such signals (sine waves that are amplitude-modulated by a Gaussian function) are called Gabor signals. Echolocation clicks are commonly of Gabor shape (e.g., Kamminga and Beitsma 1990; Holland et al. 2004). In several species of beaked whales, the sine wave is frequency-modulated (Baumann-Pickering et al. 2013) as in the example in Fig. 4.4, where the frequency changes linearly with time, sweeping up from 10 to 50 kHz.

The peak-to-peak sound pressure (symbol: $p_{pk\text{-}pk}$; unit: Pa) is the difference between the maximum pressure and the minimum pressure of a sound wave:

$$p_{pk-pk} = \max(p(t)) - \min(p(t))$$

In other words, it is the sum of the greatest magnitude during compression and the greatest magnitude during rarefaction.

The peak sound pressure (symbol: $p_{pk}$; unit: Pa) is also called zero-to-peak sound pressure and is the greatest deviation of the sound pressure from the static pressure; it is the greatest magnitude of $p(t)$:

$$p_{pk} = \max(|p(t)|)$$

This can occur during compression and/or rarefaction. In other words, $p_{pk}$ is the greater of the greatest magnitude during compression and the greatest magnitude during rarefaction (Fig. 4.1).

The root-mean-square (rms) is a useful measure for signals (like sound pressure) that aren't simple oscillatory functions. The rms of any signal can be calculated, no matter how complicated it is. To do so, square each sample of the signal, average all the squared samples, and then take the square root of the result. It turns out that the rms of a sine wave is 0.707 times its amplitude, but this is only true for sinusoidal (sine or cosine) waves. The units for rms are the same as those for amplitude (e.g., Pa if the signal is pressure or m/s if the signal is particle velocity). The root-mean-square sound pressure (symbol: $p_{rms}$; unit: Pa) is computed as its name dictates, as the root of the mean over time of the squared pressure:

$$p_{rms} = \sqrt{\frac{\int_{t_1}^{t_2} p^2(t)\mathrm{d}t}{t_2 - t_1}}, \text{or in discrete form :}$$

$$p_{rms} = \sqrt{\frac{\sum_{i=1}^{N} p_i^2}{N}} \tag{4.1}$$

This computation is practically carried out over a time interval from $t_1$ to $t_2$.

The mean-square is the mean of the square of the signal values. The mean-square of a signal is always equal to the square of the signal's rms. Its units are the square of the corresponding amplitude units (e.g., $Pa^2$ if the signal is pressure or $(m/s)^2$ if the signal is particle velocity). The mean-square sound pressure formula is similar to (Eq. 4.1) but without the root.

The sound pressure level (abbreviation: SPL; symbol: $L_p$) is the level of the root-mean-square sound pressure and computed as

$$L_p = 20 \log_{10}\left(\frac{p_{rms}}{p_0}\right)$$

expressed in dB relative to (abbreviated: re) a reference value $p_0$. The standard reference value is 20 μPa in air and 1 μPa in water.

The peak sound pressure level (also called zero-to-peak sound pressure level; abbreviation: SPLpk; symbol: $L_{p,pk}$) is the level of the peak sound pressure and computed as

$$L_{p,pk} = 20 \log_{10}\left(\frac{p_{pk}}{p_0}\right)$$

It is expressed in dB relative to a reference value $p_0$ (i.e., 20 μPa in air and 1 μPa in water). Similarly, the peak-to-peak sound pressure level is the level of the peak-to-peak sound pressure:

$$L_{p,pk-pk} = 20 \log_{10}\left(\frac{p_{pk-pk}}{p_0}\right)$$

Example sound pressure levels in air and water are given in Tables 4.3 and 4.4. Sources can have a large range of levels and only one example is given for each source. Animal sounds and their levels may vary with species, sex, age, behavioral context, etc. Animals in captivity may produce lower levels than animals in the wild. Ship noise depends on the type of vessel, its propulsion system, speed, load, etc. The tables are intended to give an overview of the dynamic range of source levels across the different sources.

Loudness is an attribute of auditory sensation. While it is related to sound pressure, loudness measures how loud or soft a sound seems to us. Given that very little is known about auditory perception in animals, the term loudness is rarely used in animal bioacoustics.

**Table 4.3** Examples of sound pressure levels in air. All levels are broadband; the hearing thresholds are single-frequency. Nominal ranges from the source are given in meters. Note that the different sources listed can have a range of levels and only one example is given

| | Pa | dB re 20 μPa |
|---|---|---|
| Explosion at 1 m | 63,246 | 190 |
| Airplane take-off at 25 m | 632 | 150 |
| Human pain threshold at 1 kHz | 200 | 140 |
| Lion roar at 1 m | 13 | 116 |
| Human discomfort threshold at 1 kHz | 10 | 114 |
| Diesel lawn mower at 1 m | 1 | 94 |
| Truck at city speed at 20 m | 0.2 | 80 |
| Old vacuum cleaner at 1 m | 0.1 | 70 |
| Bird song at 1 m | 0.02 | 60 |
| Cricket chorus at 1 m | 0.02 | 60 |
| Human speech at 1 m | 0.01 | 55 |
| Buzzing mosquito | 0.002 | 40 |
| Human whisper at 1 m | 0.001 | 30 |
| Fluttering leaves | 0.0002 | 20 |
| Human breathing at 1 m | 0.0001 | 10 |
| Human hearing threshold at 1 kHz | 0.00002 | 0 |

**Table 4.4** Examples of sound pressure levels in water. All levels are broadband; the hearing thresholds are single-frequency. Nominal ranges from the source are given in meters. Note that the different sources listed can have a range of levels and only one example is given

| | Pa | dB re 1 μPa |
|---|---|---|
| Subsea earthquake | 316,228 | 230 |
| Seismic survey airgun at 1 m | 10,000 | 200 |
| Container ship at 1 m | 5623 | 195 |
| Humpback whale song at 1 m | 1778 | 185 |
| Zodiac at high speed at 1 m | 178 | 165 |
| Dolphin whistle at 1 m | 32 | 150 |
| Geotechnical drilling at 1 m | 18 | 145 |
| Jet ski | 10 | 140 |
| Toadfish at 1 m | 10 | 140 |
| Damsel fish at 1 m | 1 | 120 |
| Open ocean ambient noise at sea state 4 | 0.1 | 100 |
| Open ocean ambient noise at sea state 0.5 | 0.01 | 80 |
| California Sea lion hearing threshold at 10 kHz | 0.001 | 60 |
| Killer whale hearing threshold at 20 kHz | 0.0001 | 40 |

### 4.2.5 Sound Exposure

Sound exposure (symbol: $E_{p,T}$; unit: $Pa^2s$) is the integral over time of the squared pressure:

$$E_{p,T} = \int_{t_1}^{t_2} p^2(t)\mathrm{d}t$$

Sound exposure increases with time. The longer the sound lasts, the greater the exposure. The sound exposure level (abbreviation: SEL; symbol: $L_{E,p}$) is computed as:

$$L_{E,p} = 10 \log_{10}\left(\frac{E_{p,T}}{E_{p,0}}\right)$$

It is expressed in dB relative to $E_{p,0} = 400$ $\mu Pa^2s$ in air, and $E_{p,0} = 1$ $\mu Pa^2s$ in water. Sound exposure is proportional to the total energy of a sound wave.

### 4.2.6 When to Use SPL and SEL?

Sound pressure and sound exposure are closely related, and in fact, the sound exposure level can be computed from the sound pressure level as:

$$L_{E,p} = L_p + 10\ \log_{10}(t_2 - t_1)$$

Conceptually, the difference is that the SPL is a time-average and therefore useful for sounds that don't change significantly over time, or that last for a long time, or that, for the assessments of noise impacts, can be considered continuous. Examples are workplace noise or ship noise. The SEL, however, increases with time and critically depends on the time window over which it is computed. It is therefore most useful for short-duration, transient sounds, such as pulses from explosions, pile driving, or seismic surveys. The SEL is then computed over the duration of the pulse.

It can be difficult to determine the actual pulse length as the exact start and end points are often not clearly visible, in particular in background noise. Therefore, in praxis, SEL is commonly computed over the 90% energy signal duration. This is the time during which 90% of the sound exposure occurs. Sound exposure is computed symmetrically about the 50% mark; i.e., from the 5% to the 95% points on the cumulative squared-pressure curve. SEL becomes (Fig. 4.5):

$$L_{E,p} = 10\ \log_{10}\left(\frac{\int_{t_{5\%}}^{t_{95\%}} p^2(\mathrm{t})\mathrm{d}t}{E_{p,0}}\right)$$

In the presence of significant background noise $p_n(t)$, the noise exposure needs to be subtracted from the overall sound exposure in order to yield the sound exposure due to the signal alone. In praxis, the noise exposure is computed over an equally long time window (from $t_1$ to $t_2$) preceding or succeeding the signal of interest:

$$L_{E,p} = 10\ \log_{10}\left(\frac{\int_{t_{5\%}}^{t_{95\%}} p^2(t)\mathrm{d}t - \int_{t_1}^{t_2} p_n^2(t)\mathrm{d}t}{E_{p,0}}\right)$$

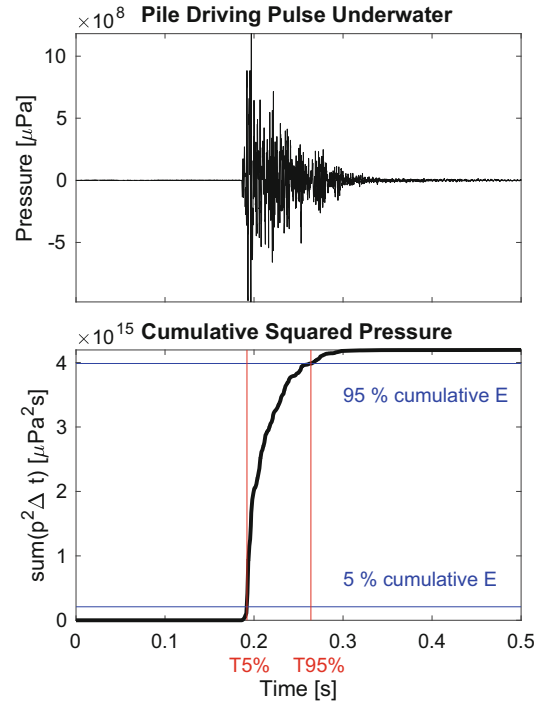



**Fig. 4.5** Pressure pulse recorded from pile driving under water (top) and cumulative squared-pressure curve (bottom). The horizontal lines indicate the 5% and 95% cumulative squared-pressure points on the y-axis. The vertical lines identify the corresponding times on the x-axis. The time between the 5% and 95% marks is the 90% energy signal duration. Recording from Erbe 2009

### 4.2.7 Acoustic Energy, Intensity, and Power

Apart from sound pressure and sound exposure, other physical quantities appear in the bioacoustics literature, but are often wrongly used. Acoustic energy refers to the total energy contained in an acoustic wave. This is the sum of kinetic energy (contained in the movement of the particles of the medium) and potential energy (i.e., work done by elastic forces in the medium). Acoustic energy $E$ is proportional to squared pressure $p$ and time interval $\Delta t$ (i.e., to sound exposure) only in the case of a free plane wave or a spherical wave at a large distance from its source:

$$E = \frac{S}{Z} p^2 \Delta t$$

The proportionality constant is the ratio of surface area $S$ through which the energy flows and acoustic impedance $Z$. Acoustic energy increases with time; i.e., the longer the sound lasts or the longer it is measured, the greater the transmitted energy. The unit of energy is joule [J] in honor of English physicist James Prescott Joule. In SI units:

$$1\ \mathrm{J} = 1\ \mathrm{kg\ m^2/s^2}$$

Acoustic power $P$ is the amount of acoustic energy $E$ radiated within a time interval $\Delta t$:

$$P = E/\Delta t$$

The unit of power is watt [W]. In SI units:

$$1\ \mathrm{W} = 1\ \mathrm{J/s} = 1\ \mathrm{kg\ m^2/s^3}$$

Acoustic intensity $I$ is the amount of acoustic energy $E$ flowing through a surface area $S$ perpendicular to the direction of propagation, per time $\Delta t$:

$$I = E/(S\Delta t) = P/S$$

For a free plane wave or a spherical wave at a large distance from its source, this becomes:

$$I = p^2/Z \tag{4.2}$$

The unit of intensity is $\mathrm{W/m^2}$. A conceptually different definition equates the instantaneous acoustic intensity with the product of sound pressure and particle velocity $u$:

$$I(t) = p(t)\ u(t)$$

The two concepts are mathematically equivalent for free plane and spherical waves and the unit of intensity is always $\mathrm{W/m^2}$.

The above quantities (energy, power, and intensity) are sometimes used interchangeably. That's wrong. They are not the same, but they are related. With $E$, $P$, $I$, $S$, and $t$ denoting energy, power, intensity, surface area, and time, respectively:

$$P = E/\Delta t = I\ S$$

More information and definitions can be found in acoustic standards (including American National Standards Institute 2013; International Organization for Standardization 2017).

### 4.2.8 Particle Velocity

Particle velocity (symbol: $u$; unit: m/s) refers to the oscillatory movement of the particles of the acoustic medium (i.e., molecules in air and water, and atoms in the ground) as a wave passes through. In the example of Fig. 4.1, the particle velocity is a sine wave, just like the acoustic pressure. Each particle oscillates about its equilibrium position. At this point, its displacement is zero, but its velocity is greatest (i.e., either maximally positive or maximally negative, depending on the direction in which the particle is moving). At the two turning points, the displacement from the equilibrium position is maximum and the velocity passes through zero, changing sign (i.e., direction) from positive to negative, or vice versa. Velocity is a vector, which means it has both magnitude and direction. Particle displacement (unit: m) and particle acceleration (unit: $\mathrm{m/s^2}$) are also vector quantities. In fact, particle velocity is the first derivative of particle displacement with respect to time, and particle acceleration is the second derivative of particle displacement with respect to time. Measurements of particle displacement, velocity, and acceleration created by snorkeling are shown in Fig. 4.6.

Air molecules also move due to wind, and water molecules move due to waves and currents. But these types of movement are not due to sound. Wind velocity and current velocity are entirely different from the oscillatory particle velocity involved in the propagation of sound.

It is equally important to understand that the speed at which the particles move when a sound wave passes through is not equal to the speed of sound at which the sound wave travels through the medium. The latter is not an oscillatory quantity.

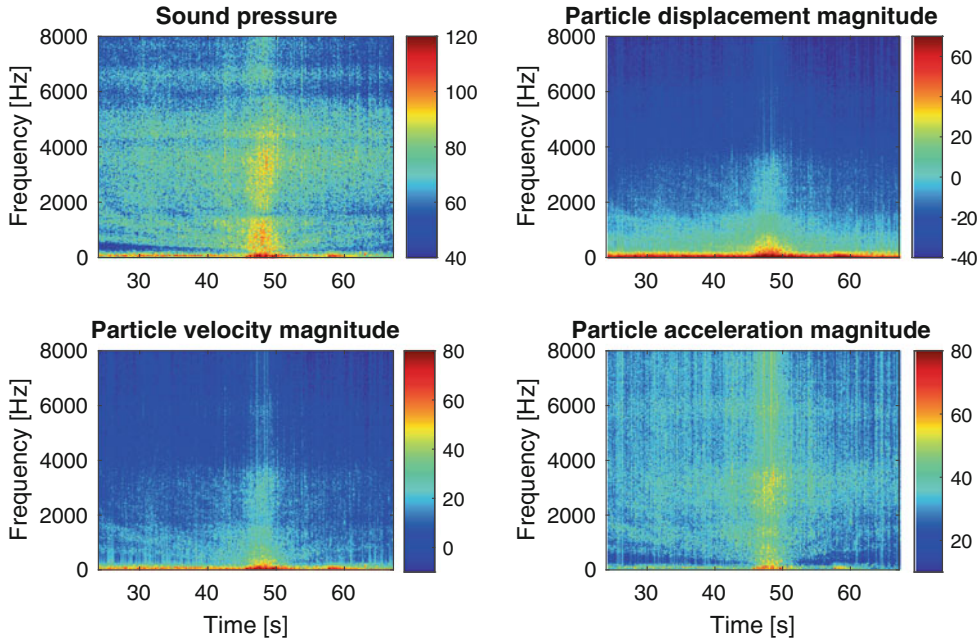



**Fig. 4.6** Spectrograms of mean-square sound pressure spectral density [dB re 1 μPa$^2$/Hz], mean-square particle displacement spectral density [dB re 1 pm$^2$/Hz], mean-square particle velocity spectral density [dB re 1 (nm/s)$^2$/Hz], and mean-square particle acceleration spectral density [dB re 1 (μm/s$^2$)$^2$/Hz] recorded under water when a snorkeler swam above the recorder (Erbe et al. 2016b; Erbe et al. 2017a)

### 4.2.9 Speed of Sound

The speed at which sound travels through an acoustic medium is called the speed of sound (symbol: $c$; unit: m/s). It depends primarily on temperature and height above ground in air, and on temperature, salinity, and depth below the sea surface in water. The speed of sound is computed as the distance sound travels divided by time. It can also be computed from measurements of the waveform (i.e., wavelength, period, and frequency as in Fig. 4.1):

$$c = \lambda/\tau = \lambda f$$

In solid media, such as rock, two types of waves are supported, P- and S-waves (see Sect. 4.2.2), and the speeds ($c_P$ and $c_S$) at which they travel differ. Table 4.5 gives examples for the speed of sound in air and water, and for P- and S-waves in some Earth materials. Example sound speed profiles (i.e., line graphs of sound speed versus altitude or water depth) are given in Fig. 4.7.

### 4.2.10 Acoustic Impedance

Each acoustic medium has a characteristic impedance (symbol: $Z$). It is the product of the medium's density (symbol: $\rho$) and speed of sound: $Z = \rho c$. In air at 0 °C with a density $\rho = 1.3$ kg/m$^3$ and speed of sound $c = 330$ m/s, the characteristic impedance is $Z = 429$ kg/(m$^2$s). In freshwater at 5 °C with a density of $\rho = 1000$ kg/m$^3$ and a speed of sound $c = 1427$ m/s, the characteristic impedance is $Z = 1427{,}000$ kg/(m$^2$s). In sea water at 20 °C and 1 m depth with 3.4% salinity, a density of $\rho = 1035$ kg/m$^3$, and a speed of sound of $c = 1520$ m/s, the characteristic impedance is $Z = 1{,}573{,}200$ kg/(m$^2$s). The characteristic impedance relates the sound pressure to particle velocity via $p = Z\,u$ for plane waves.

**Table 4.5** P-wave and S-wave speeds of certain acoustic media

| Medium | $c_P$ [m/s] | $c_S$ [m/s] |
|---|---|---|
| Air, 0 °C | 330 | |
| Air, 20 °C | 343 | |
| Freshwater, 5 °C | 1427 | |
| Freshwater, 20 °C | 1481 | |
| Salt water, 20 °C, salinity 3.4%, 1 m depth | 1520 | |
| Sand | 800–2200 | |
| Clay | 1000–2500 | |
| Sandstone | 1400–4300 | 700–2800 |
| Granite | 5500–5900 | 2800–3000 |
| Limestone | 5500–6100 | 2800–3300 |

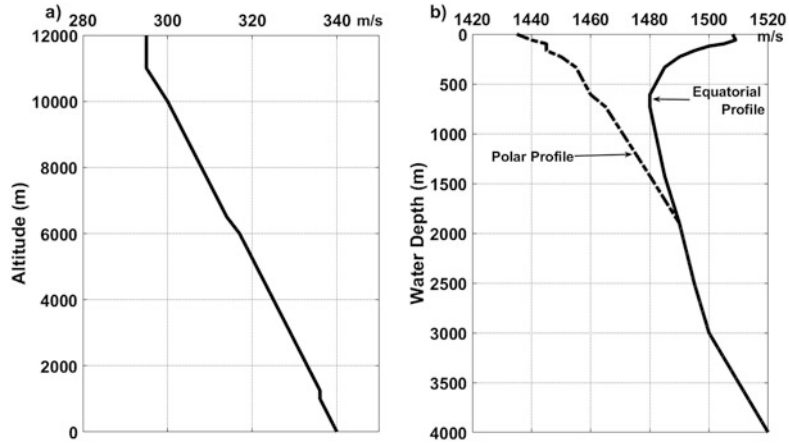



**Fig. 4.7** Example profiles of the speed of sound in (**a**) air (data from The Engineering ToolBox; https://www.engineeringtoolbox.com/elevation-speed-sound-air-d_1534.html; accessed 16 April 2021) and (**b**) water in polar and equatorial regions (These data were collected and made freely available by the International Argo Program and the national programs that contribute to it; https://argo.ucsd.edu, https://www.ocean-ops.org. The Argo Program is part of the Global Ocean Observing System. Argo float data and metadata from Global Data Assembly Centre (Argo GDAC); https://doi.org/10.17882/42182; accessed 16 April 2021). See Chaps. 5 and 6

### 4.2.11 The Decibel

Acousticians may deal with very-high-amplitude signals and very-low-amplitude signals; e.g., the sound pressure near an explosion might be 60,000 Pa, while the sound pressure from human breathing is only 0.0001 Pa. This means that the dynamic range of quantities in acoustics is large and, in fact, covers seven orders of magnitude (see Tables 4.3 and 4.4). Rather than handling multiple zeros and decimals, using a logarithmic scale compresses the dynamic range into a manageable range of values. This is one of the reasons why the decibel is so popular in acoustics. Another reason is that human perception of the loudness of a sound is approximately proportional to the logarithm of its amplitude.

When quantities such as sound pressure or sound exposure are converted to logarithmic scale, the word "level" is added to the name. Sound pressure level and sound exposure level are much more commonly used than their linear counterparts, sound pressure and sound exposure.

By definition, the level $L_Q$ of quantity $Q$ is proportional to the logarithm of the ratio of $Q$ and a reference value $Q_0$, which has the same unit. In the case of a field quantity $F$, such as sound pressure or particle velocity, or an electrical quantity such as voltage or current, the level $L_F$ is computed as

$$L_F = 20 \log_{10} \frac{F}{F_0}$$

In the case of a power quantity $P$, such as mean-square sound pressure or energy, the level $L_P$ is computed as

$$L_P = 10 \log_{10} \frac{P}{P_0}$$

Both levels are expressed in decibels (dB). Note the different factors (20 versus 10) in the equations. It is critically important to always state the reference value $F_0$ or $P_0$ when discussing levels, because reference values differ between air and water.

#### 4.2.11.1 Conversion from Decibel to Field or Power Quantities

The relationships for calculating field and power quantities from their levels are, respectively:

$$F = 10^{\frac{L_F}{20}} F_0, \text{and } P = 10^{\frac{L_P}{10}} P_0 \qquad (4.3)$$

The units of the calculated quantities correspond to the units of the reference quantity ($F_0$ or $P_0$). For example, an underwater tone at a level of 120 dB re 1 μPa rms has an rms pressure of 1 Pa. This is worked out as follows:

$$F = 10^{120/20} \times 1\mu\text{Pa} = 10^6\ \mu\text{Pa} = 1\ \text{Pa}$$

However, a tone of 120 dB re 20 μPa rms in air has an rms pressure of 20 Pa:

$$F = 10^{120/20} \times 20\ \mu\text{Pa} = 10^6 \cdot 20\ \mu\text{Pa} = 20\ \text{Pa}$$

#### 4.2.11.2 Differences between Levels of like Quantities

A particular difference between two levels corresponds to particular ratios between their field and power quantities. The general relationships are:

$$L_{F1} - L_{F2} = 20 \log_{10} \frac{F_1}{F_2}$$

$$L_{P1} - L_{P2} = 10 \log_{10} \frac{P_1}{P_2}$$

$$\frac{F_1}{F_2} = 10^{\left(\frac{L_{F1} - L_{F2}}{20}\right)}$$

$$\frac{P_1}{P_2} = 10^{\left(\frac{L_{P1} - L_{P2}}{10}\right)}$$

Some common examples are given in Table 4.6. Note the inverse relationship between ratios for corresponding positive and negative level differences and also that each power

**Table 4.6** Level differences and their corresponding field and power quantity ratios

| Level difference ($L_{F1}$-$L_{F2}$ or $L_{P1}$-$L_{P2}$) in dB | Field quantity ratio ($F_1/F_2$); use for pressure, particle velocity, voltage, current, etc. | Power quantity ratio ($P_1/P_2$); use for power, intensity, energy, sound exposure, mean-square pressure, etc. |
|---|---|---|
| −40 | 1/100 = 0.01 | 1/10,000 = 0.0001 |
| −20 | 1/10 = 0.1 | 1/100 = 0.01 |
| −10 | $1/\sqrt{10} \approx 0.316$ | 1/10 = 0.1 |
| −6 | 1/2 = 0.5 | 1/4 = 0.25 |
| −3 | $1/\sqrt{2} \approx 0.707$ | 1/2 = 0.5 |
| 0 | 1 | 1 |
| 3 | $\sqrt{2} \approx 1.41$ | 2 |
| 6 | 2 | 4 |
| 10 | $\sqrt{10} \approx 3.16$ | 10 |
| 20 | 10 | 100 |
| 40 | 100 | 10,000 |

quantity ratio is the square of the corresponding field quantity ratio.

For example, a tone at a level of 120 dB re 1 μPa rms is 20 dB stronger than a tone at a level of 100 dB re 1 μPa rms, so from Table 4.6, the ratio of the two rms pressures is $p_1/p_2 = F_1/F_2 = 10$, and the ratio of their intensities is $I_1/I_2 = P_1/P_2 = 100$.

#### 4.2.11.3 Amplification of Signals

The above formulae and Table 4.6 can also be used to calculate the effect of amplifying signals. For example, if an amplifier has a gain of 20 dB, then the rms voltage at the output of the amplifier will be 10 times the rms voltage at its input. Similarly, an amplifier with a 40 dB gain will increase the rms voltage by a factor of 100. If several amplifier stages are cascaded, then their combined gain is the sum of the gains of the individual stages (in dB).

When calibrating acoustic recordings (see Chap. 2), the gains of all components of the recording systems have to be summed. An underwater recording system (Fig. 4.8), for example, contains a hydrophone that converts received acoustic pressure to a time series of voltages at its output. The sensitivity of the hydrophone specifies this relationship. For example, a hydrophone with a sensitivity $N_S = -180$ dB re 1 V/μPa produces $10^{-180/20} = 10^{-9}$ Volts output per 1 μPa input. A more sensitive hydrophone has a less negative sensitivity. The output voltage might be passed to an amplifier with $\Delta L_G = 20$ dB gain, after which it is digitized by a data acquisition board, such as a computer's soundcard. All analog-to-digital converters have a digitization gain expressed in dB re FS/V, which specifies the input voltage that leads to full scale (FS). If the digitizer has a digitization gain $\Delta L_{DG} = 10$ dB re FS/V, then $10^{10/20}$ FS/V $= 10^{1/2}$ FS/V is the relationship between FS and input voltage, meaning that FS is reached when the input is $1/10^{1/2}$ V $= 0.32$ V. The actual value of FS depends on the number of bits available. A 16-bit digitizer in bipolar mode (i.e., producing both positive and negative numbers) has a full-scale value of $2^{16-1} = 2^{15} = 32{,}768$. And so the digital values $v$ representing the acoustic pressure will lie between −32,768 and + 32,767 (with one of the possible numbers being 0). The final steps in relating these digital values to the recorded acoustic pressure entail dividing by FS, converting to dB, and subtracting all the gains:

$$\begin{aligned} L_p &= 20\ \log_{10}(v/\mathrm{FS}) - \Delta L_{DG} - \Delta L_G - N_S \\ &= 20\ \log_{10}(v/\mathrm{FS}) + 150\ \text{dB re 1 μPa} \end{aligned}$$

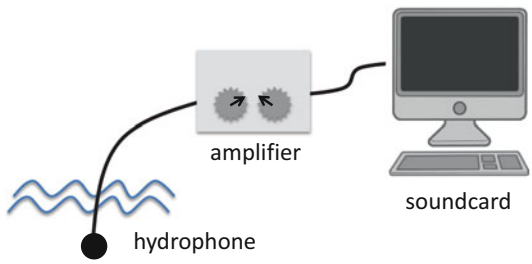



**Fig. 4.8** Sketch of an example underwater recording setup. A terrestrial setup would have a microphone instead of a hydrophone

#### 4.2.11.4 Superposition of Field and Power Quantities

If two tones of the same frequency and level arrive in phase at a listener, then the amplitude is doubled and the combined level is therefore 6 dB above the level of each tone (see Table 4.6). If, on the other hand, there is a random phase difference between the two tones then, on average, the intensity of the two signals will sum. In this case (again from Table 4.6) the combined intensity is 3 dB higher than the level of each tone. For example, if each tone has a level of 120 dB re 1 μPa rms, then the two tones together have a level of 126 dB re 1 μPa rms if they are in phase. Their superposition has an average level of 123 dB re 1 μPa rms if they have a random phase difference. Summing signals that have the same phase, or a fixed phase difference, is known as coherent summation, whereas performing an "on average" summation of signals assuming a random phase is called incoherent summation.

The calculation is more complicated if the two tones have different levels. It is necessary to use Eq. (4.3) to convert both levels to corresponding

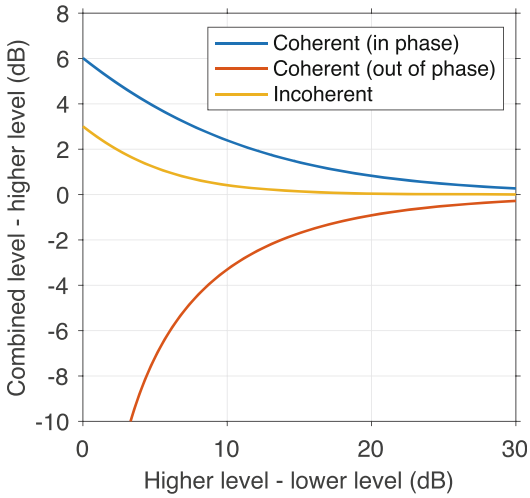



**Fig. 4.9** Line graphs of the effect on the higher-level signal of combining two signals by coherent summation (assuming the signals are in phase or 180° out of phase) and incoherent summation

field (coherent summation) or power (incoherent summation) quantities, add these quantities, and then convert the result back to a level.

The outcome of this process is plotted in Fig. 4.9 in terms of the increase in the combined level from that of the higher-level signal as a function of the difference between the higher and lower levels. Note that this increase never exceeds 6 dB for a coherent summation or 3 dB for an incoherent summation. In the case of a coherent summation, proper account has to be taken of the relative phases of the two tones when adding the field quantities, and this can have a very large effect. Figure 4.9 shows the extreme cases: The upper limit occurs when the two signals are in phase, and the lower limit occurs when they have a phase difference of 180° ($\pi$ radians). The latter case gives destructive interference and the combined level is $-\infty$!

Another useful observation from Fig. 4.9 is that when the difference in level between the two individual signals is greater than 10 dB, the incoherent summation is less than 0.5 dB higher than that of the higher of the two; and for many practical applications, the lower-level signal can be ignored.

#### 4.2.11.5 Levels in Air Versus Water

Comparing sound levels in air and water is complicated and has caused much confusion in the past. For two sound sources of equal intensity $I_a$ and $I_w$ in air and water, respectively, the sound pressure level is 62 dB greater in water because of two factors: the greater acoustic impedance of water and the different reference pressures used in the two media.

The effect of the acoustic impedance can be seen as follows. Assuming $I_w = I_a$, then from (Eq. 4.2):

$$\frac{p_w^2}{Z_w} = \frac{p_a^2}{Z_a}, \text{which is equivalent to } \frac{p_w^2}{p_a^2} = \frac{Z_w}{Z_a}.$$

This ratio of mean-square pressures in the two media can be expressed in terms of the density and speed of sound of the two media:

$$\frac{p_w^2}{p_a^2} = \frac{Z_w}{Z_a} = \frac{\rho_w c_w}{\rho_a c_a}.$$

Applying $10 \log_{10}()$ to these ratios, the difference between the mean-square sound pressure levels in water and air is:

$$\begin{aligned} L_{pw^2} - L_{pa^2} &= 10 \log_{10} \frac{p_w^2}{p_0^2} - 10 \log_{10} \frac{p_a^2}{p_0^2} \\ &= 10 \log_{10} \frac{p_w^2}{p_a^2} = 10 \log_{10} \frac{\rho_w c_w}{\rho_a c_a} \\ &= 36 \text{ dB} \end{aligned}$$

The difference between the sound pressure levels is, of course, also 36 dB:

$$\begin{aligned} L_{pw} - L_{pa} &= 20 \log_{10} \frac{p_w}{p_0} - 20 \log_{10} \frac{p_a}{p_0} \\ &= 20 \log_{10} \frac{p_w}{p_a} = 20 \log_{10} \sqrt{\frac{\rho_w c_w}{\rho_a c_a}} \\ &= 36 \text{ dB} \end{aligned}$$

In the above two equations, the same reference pressure $p_0$ is required. However, the convention

is to use $p_{a0}$=20 μPa in air and $p_{w0}$=1 μPa in water. The difference in reference pressures adds another 26 dB to the sound pressure level in water, because:

$$20 \log_{10} \frac{p_{a0}}{p_{w0}} = 20 \log_{10} \frac{20\ \mu\text{Pa}}{1\ \mu\text{Pa}} = 26\ \text{dB}$$

So, if two sound sources emit the same intensity in air and water, then the sound pressure level in water referenced to 1 μPa is 62 dB (i.e., 36 dB + 26 dB) greater than the sound pressure level in air referenced to 20 μPa.

While this might be confusing, there would hardly be a sensible reason to compare levels in air and water. Such comparisons have been attempted in the past to give an analogy to levels with which humans have experience in air. For example, humans find 114 dB re 20 μPa annoying and 140 dB re 20 μPa painful, so what would be a similarly annoying level under water that might disturb animals?

But animals perceive sound differently from humans, hear sound at different frequencies and levels, and can have rather different auditory anatomy (see Chap. 10 on audiograms). As a result, a signal easily heard by a human could be barely audible to some animals or much louder to others. Even for divers, sound reception under water is quite a different process from sound reception in air, due to different acoustic impedance ratios of the acoustic medium and human tissues, and different sound propagation paths. Furthermore, the psychoacoustic effects (emotional impacts) of different types of noise on animals have not been examined thoroughly. Even in humans, for example, 110 dB re 20 μPa of rock music does not provide the same experience as 110 dB re 20 μPa of traffic noise.

### 4.2.12 Source Level

The source level (abbreviation: SL; symbol: $L_S$) is meant to be characteristic of the sound source and independent of both the environment in which the source operates and the method by which the source level is determined. In praxis, the determination of the source level has numerous problems. Some sources are large in their physical dimensions and placing a recorder at short range (i.e., into the so-called near-field, see Sect. 4.2.13) will not result in a level that captures the full output of the source. Also, many sound sources do not operate in a free-field but rather near a boundary (e.g., air-ground, air-water, or water-seafloor). At such boundaries, reflection, scattering, absorption, and phase changes may occur, affecting the recorded level. In praxis, a sound source is recorded at some range in the far-field and an appropriate (and sometimes sophisticated) sound propagation model is utilized to account for the effects of the environment in order to compute a source level that is independent of the environment. Such source levels can then be applied to new situations and different environments in order to predict received levels elsewhere. Like other levels, the source level is expressed in dB relative to a reference value. It is further referenced to a nominal distance of 1 m from the source. The source level can be a sound pressure level or a sound exposure level, depending on the source and situation.

The radiated noise level (abbreviation: RNL; symbol $L_{RN}$) is more easily determined. It is the level of the product of the sound pressure and the range $r$ at which the sound pressure is recorded, and it can be calculated as the received sound pressure level $L_p$ plus a spherical propagation loss term:

$$L_{RN} = 20\ \log_{10} \frac{p_{rms}(r) r}{p_0 r_0} = L_p + 20\ \log_{10} \frac{r}{r_0}$$

It is expressed in dB relative to a reference value of $p_0 r_0 = 20$ μPa m in air and $p_0 r_0 = 1$ μPa m in water. The radiated noise level is dependent upon the environment and is therefore also called affected source level. Note that it is very common in the bioacoustic literature to report source levels and radiated noise levels as dB re 20 μPa @ 1 m in air and dB re 1 μPa @ 1 m in water. The ISO definition is mathematically different and the notation excludes "@ 1 m" (International Organization for Standardization 2017).

While the source level can be characteristic of the source, there are many factors that affect the source level. For example, larger ships typically

have a higher source level than smaller ships. Cars going fast have a higher source level than cars going slowly. Animals can vary the amplitude of the same sound depending on the context and their motivation. Different sound types can have different source levels. Territorial defense or aggressive sounds usually have the highest source level in a species' repertoire. Mother-offspring sounds often have the lowest source level in a species' repertoire, because mother and calf are typically close together and want to avoid detection by predators.

### 4.2.13 What Field? Free-Field, Far-Field, Near-Field

While this might read like the opening of a Dr. Seuss book, it is quite important to understand these concepts. The free-field, or free sound field, exists around a sound source placed in a homogeneous and isotropic medium that is free of boundaries. Homogenous means that the medium is uniform in all of its parameters; isotropic means that the parameters do not depend on the direction of measurement. While the free-field assumption is commonly applied to estimates of particle velocity from pressure measurements or estimates of propagation loss, sound sources and receivers are rarely in a free-field. More often, sound sources and receivers are near a boundary. This is the case for sources such as trains or construction sites and for receivers such as humans, all of which are right at the air-ground boundary. This is also the case for sources such as ships at the water surface and for receivers such as fishes in shallow water, where they are near two boundaries: the air-water and the water-seafloor boundaries. At boundaries, some of the sound is transmitted into the other medium, some of it is reflected, some of it is scattered in various directions. For more detail on source-path-receiver models in air and water, see Chaps. 5 and 6.

The far-field is the region that is far enough from the source so that the particle velocity and pressure are effectively in phase. The near-field is the region closer to the source where they become out of phase either because sound from different parts of the source arrives at different times (This is the case of an extended source.) or because the curvature of the spherical wavefront from the source is too great to be ignored (This is the case of a source small enough to be considered a point source.). These two cases have different frequency dependence with the near-field to far-field transition distance increasing with increasing frequency for an extended source, and decreasing with increasing frequency for a small source. A single source may behave as a small source at low frequencies and as an extended source at high frequencies, which implies that there is some non-zero frequency at which it will have a minimum near-field to far-field transition distance. This has resulted in much confusion.

When is a sound source small versus extended? A sound source can be considered small when its physical dimensions are small compared to the acoustic wavelength. A fin whale (*Balaenoptera physalus*) with a head size of perhaps 6 m produces a characteristic 20-Hz signal that has a wavelength of about 70 m and so the whale can be considered small.

When studying the effects of noise on animals, however, the noise sources one deals with are mostly extended sources. In the near-field, the amplitudes of field and power quantities are affected by the physical dimension of the sound source. This is because the surface of an extended sound source can be considered an array of separate point sources. Each point source generates an acoustic wave. At any location, the instantaneous pressure (as an example of a field quantity) is the summation of the instantaneous pressures from all of the point sources. In the near-field, the various sound waves have traveled various distances and arrive at various phases. Therefore, the near-field consists of regions of destructive and constructive interference and the pressure amplitude depends greatly on where exactly in the near-field it is measured. There may be regions close to a sound source where the pressure amplitude is always zero. The interference pattern depends on the frequency of the sound, and the regions of destructive and constructive interference will be different depending on the

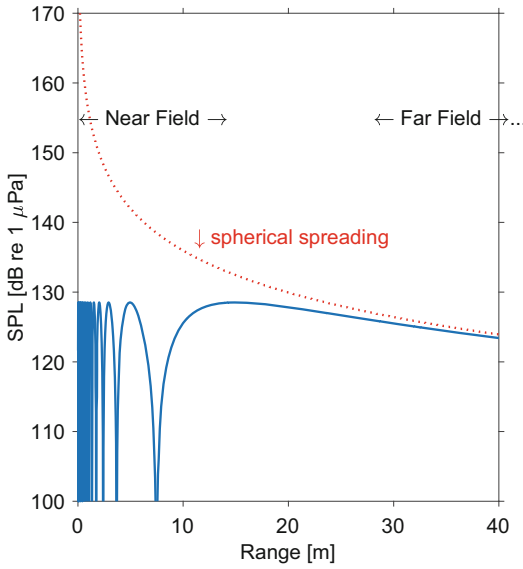



**Fig. 4.10** Graph of sound pressure versus range, perpendicular from a circular piston such as a loudspeaker with radius 1 m, f = 22 kHz, under water

frequency of the sound. In the far-field of the extended source, the sound waves from the separate point sources have traveled nearly the same distance and arrive in phase. The pressure amplitude depends only on the range from the source and decreases monotonically with increasing range. The amplitudes of field quantities $F$ and power quantities $P$ decay with range $r$ as:

$$F(r) \sim \frac{1}{r} \text{ and } P(r) \sim \frac{1}{r^2} \text{ in the far-field.}$$

The range at which the field transitions from near to far can be estimated as $L^2/\lambda$, where $L$ is the largest dimension of the source and $\lambda$ is the wavelength of interest. (Fig. 4.10).

All sound sources have near- and far-fields. The source level of a sound source is, in praxis, determined from measurements in the far-field by correcting for propagation loss. In the example of Fig. 4.10, the sound pressure level might be measured as 126 dB re 1 μPa at 30 m range from the source. A spherical propagation loss term ($20 \log_{10} \frac{r}{r_0} = 30$ dB ; red dashed line in Fig. 4.10) is then applied to estimate the radiated noise level: 156 dB re 1 μPa m. This level is higher than what would be measured with a receiver in the near-field (blue solid line in Fig. 4.10).

Radiated noise levels and source levels are useful to estimate the received level at some range in the far-field. They will always be higher than the levels that exist in the near-field. There has been a lot of confusion about this in the bioacoustics community, for example in the case of marine seismic surveys. A seismic airgun array (i.e., a number of separate seismic airguns arranged in a 2-dimensional array) might have physical dimensions of several tens of meters and a source level (in terms of sound exposure) of 220 dB re 1 μPa$^2$s m (e.g., Erbe and King 2009). However, *in situ* measurements near the array may never exceed 190 dB re 1 μPa$^2$s, except in the immediate vicinity (<< 1 m) of an individual airgun. This is because the highest level that may be recorded is close to an individual airgun in the array. The other airguns in the array are too far away to significantly add to the level of any particular airgun (see Fig. 4.9). At short range from the array, the sound waves from some airguns will add constructively and from others destructively, so that the measured pressure amplitude is always less than the amplitude from one airgun multiplied by the number of airguns in the array. Constructive superposition of sound waves from all airguns only happens in the far-field, where the pressure amplitude is reduced due to propagation loss.

### 4.2.14 Frequency Weighting

Frequency weightings are mathematical functions applied to sound measurements to compensate quantitatively for variations in the auditory sensitivity of humans and non-human animals (see Chap. 10 on audiometry). These functions "weight" the contributions of different frequencies to the overall sound level, de-emphasizing frequencies where the subject's auditory sensitivity is less and emphasizing frequencies where it is greater. Frequency weighting essentially applies a band-pass filter to the sound. Weighting is applied before the

calculation of broadband SPLs or SELs. A number of weighting functions exist for different purposes: for example, A, B, C, D, Z, FLAT, and Linear frequency weightings to measure the effect of noise on humans. However, at present, only weightings A, C, and Z are standardized (International Electrotechnical Commission 2013).

#### 4.2.14.1 A, C, and Z Frequency Weightings

A, C, and Z frequency weightings are derived from standardized equal-loudness contours. These are curves which demonstrate SPL variations over the frequency spectrum for which constant loudness is perceived (Suzuki and Takeshima 2004). Loudness is the human perception of sound pressure. Loudness levels are measured in units of phons, determined from referencing the equal-loudness contours. The number of phons $n$ is equal in intensity to a 1-kHz tone with an SPL of $n$ dB. The equal-loudness contours were developed from human loudness perception studies (Fletcher and Munson 1933; Robinson and Dadson 1956; Suzuki and Takeshima 2004) and are standardized (International Organization for Standardization 2003). Table 4.7 defines the A, C, and Z-weighting values at frequencies up to 16 kHz. Figure 4.11 displays the contours of the weightings.

A-weighting is the primary weighting function for environmental noise assessment. It covers a broad range of frequencies from 20 Hz to 20 kHz.

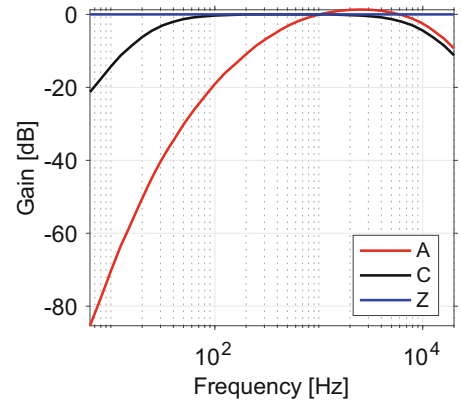



**Fig. 4.11** Graph of A-, C-, and Z-weighting curves

The function is tailored to the perception of low-level sounds and represents an idealized human 40-phon equal-loudness contour. Measurements are noted as dB(A) or dBA.

The C-weighting function provides a better representation of human auditory sensitivity to high-level sounds. This weighting is useful for stipulating peak or impact noise levels and is used for the assessment of instrument and equipment noise.

The Z-weighting function (also known as the zero-weighting function) covers a range of frequencies from 8 Hz to 20 kHz (within ± 1.5 dB), replacing the "FLAT" and "Linear" weighting functions. It adds no "weight" to account for the auditory sensitivity of humans and is commonly used in octave-band analysis to analyze the sound source rather than its effect.

**Table 4.7** A, C, and Z-weighting values

| Frequency [Hz] | A-weighting [dB] | C-weighting [dB] | Z-weighting [dB] |
|---|---|---|---|
| 63 | −26.2 | −0.8 | 0 |
| 125 | −16.1 | −0.2 | 0 |
| 250 | −8.6 | 0 | 0 |
| 500 | −3.2 | 0 | 0 |
| 1000 | 0 | 0 | 0 |
| 2000 | 1.2 | −0.2 | 0 |
| 4000 | 1 | −0.8 | 0 |
| 8000 | −1.1 | −3 | 0 |
| 16,000 | −6.6 | −8.5 | 0 |

#### 4.2.14.2 Frequency Weightings for Non-human Animals

Equal-loudness contours for non-human animals are very challenging to develop as it is difficult to obtain the required data. Direct measurements of equal loudness in non-human animals have only been achieved for bottlenose dolphins (*Tursiops truncatus*; Finneran and Schlundt 2011); however, equal-response-latency curves have been generated from reaction-time studies and been used as proxies for equal-loudness contours (Kastelein et al. 2011). Several functions applicable to the assessment of noise impact on marine mammals have also been developed similar to the A-weighting function with adjustments for the hearing sensitivity of different marine mammal groups. Other weighting functions exist for other species.

#### 4.2.14.3 M-Weighting

The M-weighting function was developed to account for the auditory sensitivity of five functional hearing groups of marine mammals (Southall et al. 2007). Development of this function was restricted by data availability and is limited in its capacity to capture all complexities of marine mammal auditory responses (Tougaard and Beedholm 2019). The function deemphasizes the frequencies near the upper and lower limits of the auditory sensitivities of each hearing group, emphasizing frequencies where exposure to high-amplitude noise is more likely to affect the focal species (Houser et al. 2017). M-weighted SEL is calculated through energy integration over all frequencies following the application of the M-weighting function to the noise spectrum. The M-weighting functions have continued to evolve, reflecting the advancement in marine mammal auditory sensitivity and response research, with the most recent modifications proposed by Southall et al. (2019), including a redefinition of marine mammal hearing groups, function assumptions, and parameters. The updated functions are based on the following equation:

$$W(f) = C + 10\log_{10}\frac{\left(\frac{f}{f_1}\right)^{2a}}{\left(\left[1+\left(\frac{f}{f_1}\right)^{2}\right]^{a}\left[1+\left(\frac{f}{f_2}\right)^{2}\right]^{b}\right)} \quad (4.4)$$

$W(f)$ is the weighting function amplitude [dB] at frequency $f$ [kHz]; $f_1$ and $f_2$ are the low-frequency and high-frequency cut-off values [kHz], respectively. Constants $a$ and $b$ are the low-frequency and high-frequency exponent values, defining the rate of decline of the weighting amplitude at low and high frequencies, and $C$ defines the vertical position of the curve (maximum weighting function amplitude is 0). Table 4.8 lists the function constants for each marine mammal hearing group and Fig. 4.12 plots the weighting curves.

### 4.2.15 Frequency Bands

Different sound sources emit sound at different frequencies and cover different frequency bands. The whistle of a bird is quite tonal, covering a narrow band of frequencies. An echosounder

**Table 4.8** Constants of Eq. 4.4 for the six functional hearing groups of marine mammals (Southall et al. 2019)

| Marine mammal hearing group | $a$ | $b$ | $f_1$ [kHz] | $f_2$ [kHz] | $C$ [dB] |
|---|---|---|---|---|---|
| Low-frequency cetaceans (LF) | 1 | 2 | 0.2 | 19 | 0.13 |
| High-frequency cetaceans (HF) | 1.6 | 2 | 8.8 | 110 | 1.20 |
| Very-high-frequency cetaceans (VHF) | 1.8 | 2 | 12 | 140 | 1.36 |
| Sirenians (SI) | 1.8 | 2 | 4.3 | 25 | 2.62 |
| Phocid carnivores in water (PCW)<br>Phocid carnivores in air (PCA) | 1<br>2 | 2<br>2 | 1.9<br>0.75 | 30<br>8.3 | 0.75<br>1.50 |
| Other marine carnivores in water (OCW)<br>Other marine carnivores in air (OCA) | 2<br>1.4 | 2<br>2 | 0.94<br>2 | 25<br>20 | 0.64<br>1.39 |

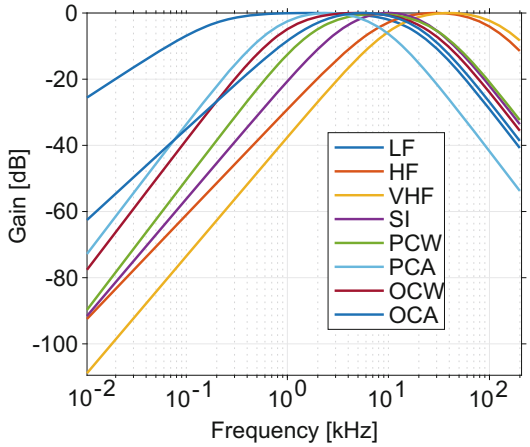



**Fig. 4.12** Weighting curves calculated from the function $W(f)$ (Eq. 4.4) and constants (Table 4.8), for each marine mammal hearing group

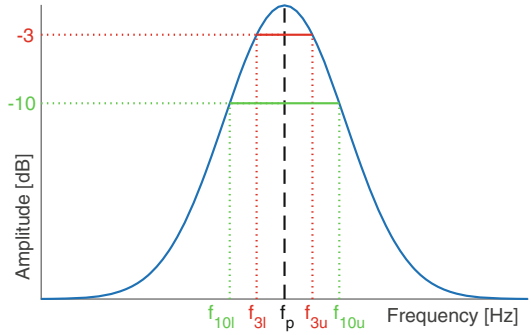



**Fig. 4.13** Illustration of the 3-dB and 10-dB bandwidths of a signal; p: peak, l: lower, u: upper

emits a sharp tone, concentrating almost all acoustic energy in a narrow frequency band centered on one frequency. These are narrowband sources, while a ship propeller is a broadband source generating many octaves in frequency. The term frequency band refers to the band of frequencies of a sound. The bandwidth is the difference between the highest and the lowest frequency of a sound. The spectrum of a sound shows which frequencies are contained in the sound and the amplitude at each frequency.

Peak frequency and 3-dB bandwidth are often used to describe the spectral characteristics of a signal. Peak frequency is the frequency of maximum power of the spectrum. The 3-dB bandwidth is computed as the difference between the frequencies (on either side of the peak frequency), at which the spectrum has dropped 3 dB from its maximum (Fig. 4.13). Remember that a drop of 3 dB is equal to half power; and so the 3-dB bandwidth is the bandwidth at the half-power marks. Similarly, the 10-dB bandwidth is measured 10 dB down from the maximum power (i.e., where the power has dropped to one tenth of its peak).

For non-Gaussian spectra (e.g., bat or dolphin echolocation clicks), two other measures are useful: the center frequency $f_c$, which splits the power spectrum into two halves of equal power, and the rms bandwidth $BW_{rms}$, which measures the standard deviation about the center frequency. With $H(f)$ representing the Fourier transform, these quantities are computed as (Fig. 4.14):

$$f_c = \frac{\int_{-\infty}^{\infty} f|H(f)|^2 \mathrm{d}f}{\int_{-\infty}^{\infty} |H(f)|^2 \mathrm{d}f}$$

$$BW_{rms} = \sqrt{\frac{\int_{-\infty}^{\infty} (f - f_c)^2 |H(f)|^2 \mathrm{d}f}{\int_{-\infty}^{\infty} |H(f)|^2 \mathrm{d}f}}$$

Broadband sounds are commonly analyzed in specific frequency bands. In other words, the energy in a broadband sound can be split into a series of frequency bands. This splitting is done by a filter, which can be implemented in hardware or software. A low-pass filter lets low frequencies pass and reduces the amplitude of (i.e., attenuates) signals above its cut-off frequency. A high-pass filter lets high frequencies pass and reduces the amplitude of signals below its cut-off frequency. A band-pass filter passes signals within its characteristic pass-band (extending from a lower edge frequency to an upper edge frequency) and attenuates signals outside of this band. It is a common misconception that a filter removes all energy beyond its cut-off frequency. Instead, a filter progressively attenuates the

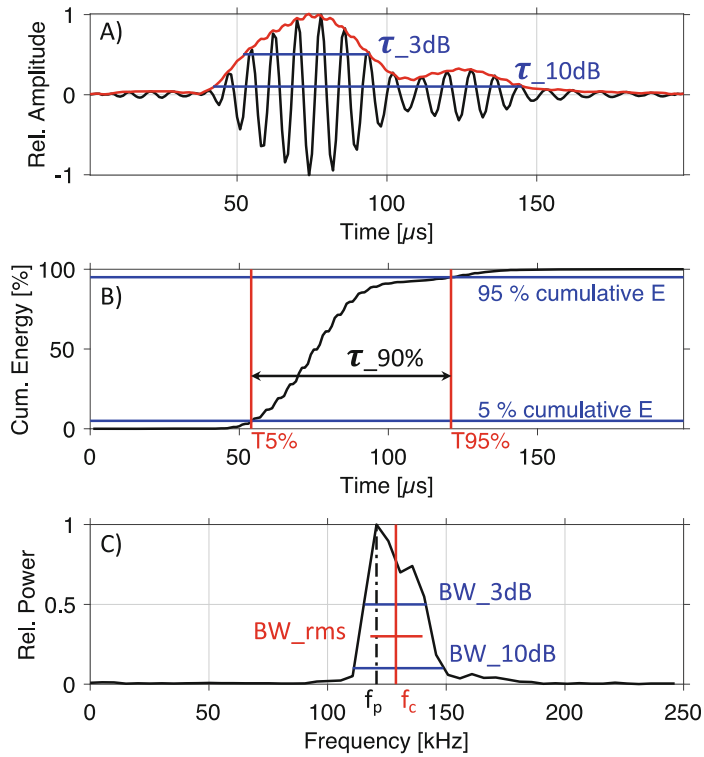



**Fig. 4.14** Echolocation click from a harbor porpoise (*Phocoena phocoena*); (**a**) waveform and amplitude envelope (determined by Hilbert transform), (**b**) cumulative energy, and (**c**) spectrum. Three different duration parameters (τ) are shown. The 3-dB duration is the difference in time between the two points at half power (i.e., 3 dB down from the maximum of the signal envelope). The 10-dB duration is the time difference between the points at one tenth of the peak power (i.e., 10 dB below the maximum). Computation of the 90% energy signal duration was explained in Sect. 4.2.6. Three bandwidth measures are shown. The 3-dB and 10-dB bandwidths are measured down from the maximum power, which occurs at the peak frequency $f_p$, and the rms bandwidth is measured about the center frequency $f_c$. Click recording courtesy of Whitlow Au

energy. At the cut-off frequency, the energy is typically reduced by 3 dB. Beyond the cut-off frequency, the attenuation increases; how rapidly depends on the order of the filter.

Band-pass filtering is very common in the study of broadband sounds, in particular broadband noise such as aircraft or ship noise. A number of band-pass filters are used that have adjacent pass-bands such that the sound spectrum is split into adjacent frequency bands. If these bands all have the same width, then the filters are said to have constant bandwidth. In contrast, proportional bandwidth filters split sound into adjacent bands that have a constant ratio of upper to lower frequency. These bands become wider with increasing frequency (e.g., octave bands).

Octave bands are exactly one octave wide, with an octave corresponding to a doubling of frequency. The upper edge frequency of an octave band is twice the lower edge frequency of the band: $f_{up} = 2\ f_{low}$. Fractional octave bands are a fraction of an octave wide. One-third octave bands are common. The center frequencies $f_c$ of adjacent 1/3 octave bands are calculated as $f_c(n) = 2^{n/3}$, where $n$ counts the 1/3 octave bands. The lower and upper frequencies of band $n$ are calculated as:

$$f_{low}(n) = 2^{-1/6}\ f_c(n) \text{ and } f_{up}(n) = 2^{1/6}\ f_c(n)$$

Another example for proportional bands are decidecades. Their center frequencies $f_c$ are

**Table 4.9** Center frequencies of adjacent 1/3 octave bands [Hz]. The table can be extended to lower and higher frequencies by division and multiplication by 10, respectively

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 10 | 12.5 | 16 | 20 | 25 | 31.5 | 40 | 50 | 63 | 80 |
| 100 | 125 | 160 | 200 | 250 | 315 | 400 | 500 | 630 | 800 |
| 1000 | 1250 | 1600 | 2000 | 2500 | 3150 | 4000 | 5000 | 6300 | 8000 |
| 10,000 | 12,500 | 16,000 | 20,000 | 25,000 | 31,500 | 40,000 | 50,000 | 63,000 | 80,000 |

calculated as $f_c(n) = 10^{n/10}$, where $n$ counts the decidecades. The lower and upper frequencies of band $n$ are calculated as:

$$f_{low}(n) = 10^{-1/20}\, f_c(n)$$
$$f_{up}(n) = 10^{1/20}\, f_c(n)$$

Decidecades are a little narrower than 1/3 octaves by about 0.08%. Decidecades are often erroneously called 1/3 octaves in the literature. Given this confusion and inconsistencies in rounding, preferred center frequencies have been published (Table 4.9).

### 4.2.16 Power Spectral Density

The spectral density of a power quantity is the average of that quantity within a specified frequency band, divided by the bandwidth of that band. Spectral densities are typically computed for mean-square sound pressure or sound exposure. Furthermore, spectral densities are most commonly computed in a series of adjacent constant-bandwidth bands, where each band is exactly 1 Hz wide. The spectral density then describes how the power quantity of a sound is distributed with frequency. The mean-square sound pressure spectral density level is expressed in dB:

$$L_{p,f} = 10\ \log_{10}\left(\frac{\overline{p_f^2}}{p_{f_0}^2}\right)$$

The reference value $p_{f_0}^2$ is 1 $\mu Pa^2/Hz$ in water. In air, it is more common to take the square root and report spectral density in dB re 20 $\mu Pa/\sqrt{Hz}$.

### 4.2.17 Band Levels

Band levels are computed over a specified frequency band. Band levels can be computed from spectral densities by integrating over frequency before converting to dB.

Consider the sketched mean-square sound pressure spectral density as a function of frequency (Fig. 4.15). The band level $L_p$ in the band from $f_{low}$ to $f_{up}$ is the total mean-square sound pressure in this band:

$$\begin{aligned} L_p &= 10\ \log_{10}\left(\frac{\int_{f_{low}}^{f_{up}} p_f^2 \mathrm{d}f}{p_{f_0}^2 f_0}\right) \\ &= 10\ \log_{10}\left(\frac{\overline{p_f^2}\left(f_{up} - f_{low}\right)}{p_{f_0}^2 f_0}\right) \\ &= 10\ \log_{10}\left(\frac{\overline{p_f^2}}{p_{f_0}^2}\right) \\ &\quad +10\ \log_{10}\left(\frac{f_{up} - f_{low}}{f_0}\right) \end{aligned}$$

where the reference frequency $f_0$ is 1 Hz. The band level of mean-square sound pressure is thus equal to the level of the average mean-square sound pressure spectral density plus 10 $\log_{10}$ of the bandwidth. The band level is expressed in dB re 1 $\mu Pa^2$ in water. In the in-air literature, it is more common to take the square root and report band levels in dB re 20 μPa. The frequency band should always be reported as well.

The wider the bands, the higher the band levels, as illustrated for 1/12, 1/3, and 1 octave bands in Fig. 4.16.

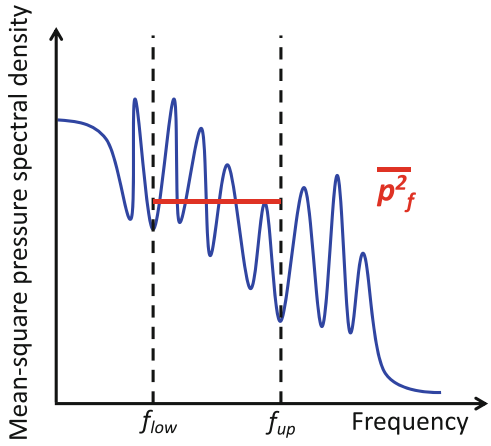



**Fig. 4.15** Graph of mean-square pressure spectral density (blue) and its average $\overline{p^2}_f$ (red) in the frequency band from $f_{low}$ to $f_{up}$

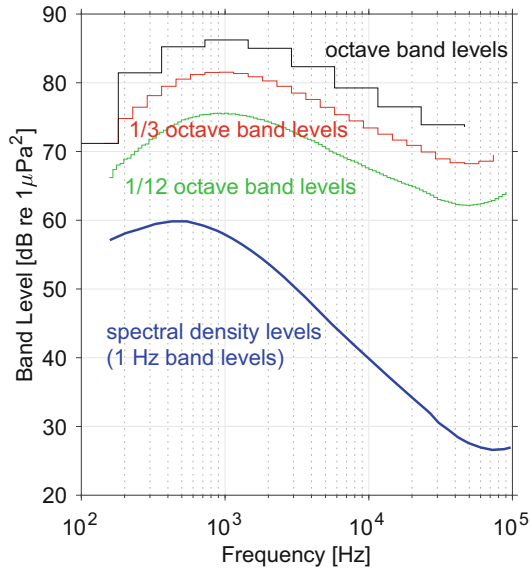



**Fig. 4.16** Illustration of band levels versus spectral density levels, for the example of wind-driven noise under water at Sea State 2. Band levels are at least as high as the underlying spectral density levels. There are twelve 1/12-octave bands in each octave, and three 1/3-octave bands. The wider the band, the higher the level, because more power gets integrated

## 4.3 Acoustic Signal Processing

### 4.3.1 Displays of Sounds

A signal can be represented in the time domain and displayed as a waveform, or in the frequency domain and displayed as a spectrum. Waveform plots typically have time on the x-axis and amplitude on the y-axis. Waveform plots are useful for analysis of short pulses or clicks. Before the common use of desktop computers, acoustic waveforms were commonly displayed by oscilloscopes (or oscillographs). The display of the waveform was called an oscillogram. Power spectra are typically displayed with frequency on the x-axis and amplitude on the y-axis.

A few examples of waveforms and their spectra are shown in Fig. 4.17.[2] A constant-wave sinusoid (a) has a spectrum consisting of a single spike at the signal's fundamental frequency, in this case 1 kHz. The signal shown in (b) has the same fundamental frequency of 1 kHz, but its spectrum shows additional overtones at integer multiples of the fundamental that are due to its more complicated shape. A pulse (c) has a quite different spectrum to the previous repetitive signals, with a maximum at zero frequency and decaying in a series of ripples (known as sidelobes) that decrease in amplitude as frequency increases. It turns out that the shorter the pulse is, the wider is the initial spectral peak. Also, the faster the rise and fall times are, the more pronounced the sidelobes are and the slower they decay. Panel (d) shows the waveform and spectrum of a 1-kHz sinusoidal signal that has been amplitude-modulated by the pulse shown in (c). The effect of this is to shift the spectrum of the pulse so that what was at zero frequency is now at the fundamental frequency of the sinusoid, and to mirror it around that frequency. Another way of thinking about this is that the effect of truncating the sinusoid is to broaden its spectrum from the spike shown in (a). The effect of changing the frequency during the burst can be seen in (e). In this case, the frequency has been swept from 500 Hz to 1500 Hz over the 10-ms burst duration. This has the effect of broadening the spectrum and smoothing out the sidelobes that were

[2] Dan Russell's animations of the Fourier compositions of different waveforms: https://www.acs.psu.edu/drussell/Demos/Fourier/Fourier.html; accessed 12 October 2020.

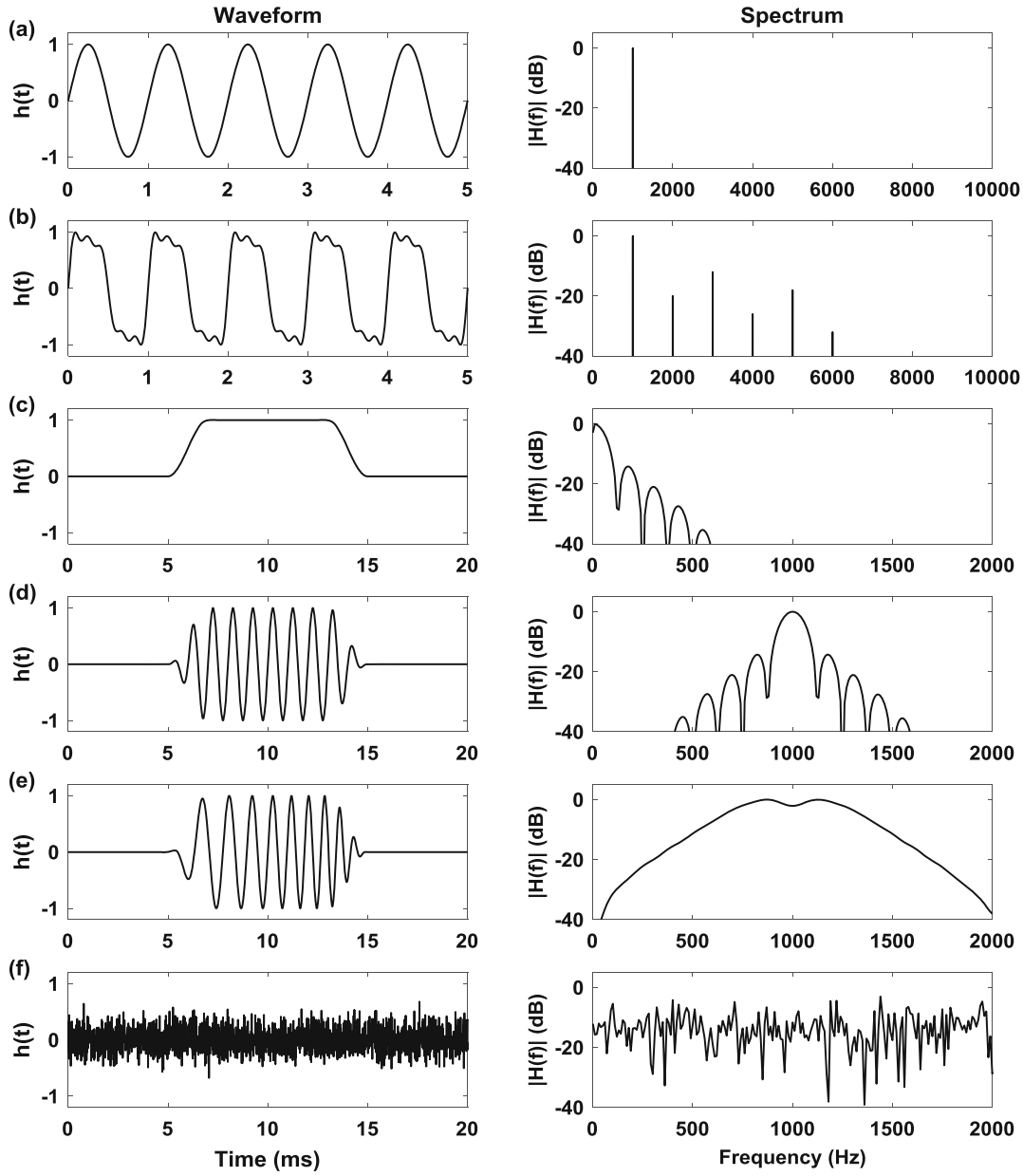



**Fig. 4.17** Examples of signal waveforms (left) and their spectra (right). (**a**) A sine wave with a frequency of 1000 Hz; (**b**) a signal consisting of a sine wave with a fundamental frequency of 1000 Hz and five overtones; (**c**) a 10-ms long pulse with 2-ms rise and fall times; (**d**) a 10-ms long tone burst with a center frequency of 1000 Hz and 2-ms rise and fall times; (**e**) a 10-ms long FM sweep from 500 Hz to 1500 Hz with 2-ms rise and fall times; and (**f**) uncorrelated (white) random noise

apparent in (d). Finally, (f) shows a waveform consisting of uncorrelated noise and its spectrum. In this context "uncorrelated" means that knowledge of the noise at one time instant gives no information about what it will be at any other time instant. This type of noise is often called white noise because it has a flat spectrum (like white light), but as can be seen in this example,

the spectrum of any particular white noise signal is itself quite noisy and it is only flat if one averages the spectra of many similar signals, or alternatively the spectra of many segments of the same signal.

A spectrogram is a plot with, most commonly, time on the x-axis and frequency on the y-axis. A quantity proportional to acoustic power is displayed by different colors or gray levels. If properly calibrated, a spectrogram will show mean-square sound pressure spectral density. A spectrogram is computed as a succession of Fourier transforms. A window is applied in the time domain containing a fixed number of samples of the digital time series. The Fourier transform is computed over these samples. Amplitudes are squared to yield power. The power spectrum is then plotted as a vertical column with frequency on the y-axis. The window in the time domain is then moved forward in time and the next samples of the digital time series are taken and Fourier-transformed. This second spectrum is then plotted next to the first spectrum, as the second vertical column in the spectrogram. The window in the time domain is moved again, the third Fourier transform is computed and plotted as the third column of the spectrogram, and so forth (see examples in Fig. 4.2). The spectrogram, therefore, shows how the spectrum of a sound changes over time. With modern signal processing software, researchers are able to listen to the sounds in real-time while viewing the spectral patterns.

### 4.3.2 Fourier Transform

It turns out that any signal can be broken down into a sum of sine waves with different amplitudes, frequencies, and phases. This is done by the Fourier transform, named after French mathematician and physicist Joseph Fourier. While the original signal can be represented as a time series $h(t)$ (e.g., sound pressure $p(t)$) in the time domain, the Fourier transform transforms the signal into the frequency domain, where it is represented as a spectrum $H(f)$. The magnitude of $H$ is the amount of that frequency in the original signal. $H(f)$ is a complex function and the argument contains the phase of that frequency. The inverse Fourier transform recreates the original signal from its Fourier components. For a continuous function with $t$ representing time and $f$ representing frequency, the Fourier transform is ($i$ is the imaginary unit):

$$H(f) = \int_{-\infty}^{\infty} h(t) e^{-2\pi i f t} \mathrm{d}t$$

and the inverse Fourier transform is:

$$h(t) = \int_{-\infty}^{\infty} H(f) e^{2\pi i f t} \mathrm{d}f$$

While a sound wave might be continuous, during digital recording or digitization of an analogue recording, its instantaneous pressure is sampled at equally spaced times over a finite window in time. This results in a finite and discrete time series. The equations for the discrete Fourier transform are similar to the above, where the integrals are replaced by summations. The fast Fourier transform (FFT) is the most common mathematical algorithm for computing the discrete Fourier transform. In animal bioacoustics, the FFT is the most commonly used algorithm to compute the frequency spectrum of a sound. The most common display of the frequency spectrum is as a power spectrum. Here, the amplitudes $H(f)$ are squared and in this process, the phase information is lost and, therefore, the original time series cannot be recreated. If sufficient care is taken to properly preserve the phase information, it is not only possible, but often very convenient, to transform a signal into the frequency domain using the FFT, carry out processing (such as filtering) in this domain, and then use an inverse FFT to resynthesize the processed signal in the time domain.

### 4.3.3 Recording and FFT Settings

Sounds in the various displays can look rather different depending on the recording and analysis parameters. There is no set of parameters that will produce the best display for all sounds. Rather,

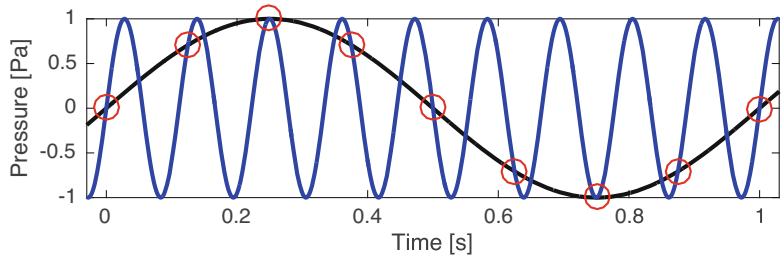



**Fig. 4.18** Waveforms of a 1-Hz sine wave (black) and a 9-Hz sine wave (blue), both sampled 8 times per second (i.e., $f_s = 8$ Hz) as indicated by the red circles. Note that the red samples fit either sine wave. In fact, there is an infinite number of signals that fit these samples

the ideal parameters depend on the question being asked, and it is important to have a thorough understanding of each of the parameters or selectable settings, and how they interact.

#### 4.3.3.1 Sampling Rate

Microphones and hydrophones produce continuous voltages in response to sounds. The voltage outputs are termed analogue in that they are direct analogues of the acoustic signal. Analogue-to-digital converters sample the voltages of the signal and the level is expressed as a number (a digit) for each of the samples. The sampling rate is the number of samples per second and its unit is 1/s. The inverse is called the sampling frequency (symbol: $f_s$; unit: Hz). Music on commercial CDs is digitized at 44.1 kHz (i.e., there are 44,100 samples stored every second). At high sampling rates, the digital sound file becomes very large for long-duration sound. The rate at which sounds are sampled by a digital recorder is typically stored in the header of the sound file. This file is a list of numbers with each number being the sound pressure at that sample point. Digital sound files are an incomplete record of the original signal; the intervals in the original signal between samples are lost during digitizing. The result is that there is a maximum frequency (related to the sampling rate) that can be resolved during Fourier analysis. Imagine a low-frequency sine wave. Only a few samples are needed to determine its frequency and amplitude and to recreate the full sine wave (by interpolation) from its samples. Those few samples might not be enough if the frequency is higher.

#### 4.3.3.2 Aliasing

Aliasing is a phenomenon that occurs due to sampling. A continuous acoustic wave is digitally recorded by sampling at a sampling frequency $f_s$ and storing the data as a time series $p(t)$. It turns out that different signals can produce the identical time series $p(t)$ and are therefore called aliases of each other. In Fig. 4.18, $p_{black}(t)$ has a frequency $f_{black} = 1$ Hz, while $p_{blue}(t)$ has a frequency $f_{blue} = 9$ Hz. A recorder that samples at $f_s = 8$ Hz would measure the pressure as indicated by the red circles from either the red or the blue time series. Based on the samples only, it is impossible to tell which was the original time series. In fact, there is an infinite number of signals that fit these samples. If $f_0$ is the lowest frequency that fits these samples, then the frequency of the $n^{\text{th}}$ alias is $f_a(n)$, with $n$ being an integer number:

$$\frac{f_a(n)}{f_s} = \frac{f_0}{f_s} + n$$

The most common problem of aliasing in animal bioacoustics occurs if a high-frequency animal sound is recorded at too low a sampling frequency. After FFT, the spectrum or spectrogram displays a sound at an erroneously low frequency. The Nyquist frequency (named after Harry Nyquist, a Swedish-born electronic engineer) is the maximum frequency that can be determined and is equal to half the sampling frequency. This requires some *a priori* information of the sounds to be recorded before a recording system is put together. The

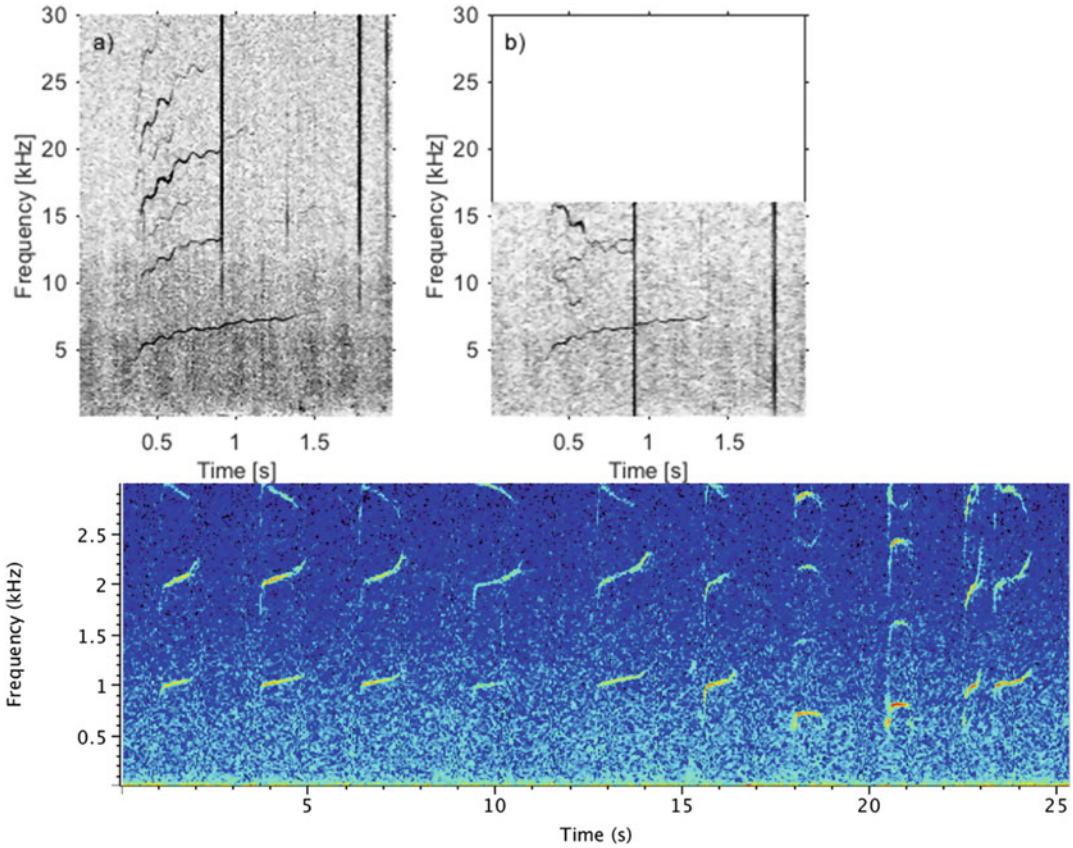



**Fig. 4.19** Examples of folding (aliasing). Top: A killer whale sound sampled at 96 kHz (**a**) and at 32 kHz (**b**) (Wellard et al. 2015). If no anti-aliasing filter is applied, frequencies above the Nyquist frequency (i.e., 16 kHz in the right panel) will appear reflected downwards; upsweeps greater than the Nyquist frequency appear as downsweeps. Bottom: Humpback whale (*Megaptera novaeangliae*) notes recorded with a sampling frequency of 6 kHz, but without an anti-aliasing filter. Contours above 3 kHz appear mirrored about the 3-kHz edge

higher the sampling frequency is, the higher the maximum frequency that can be accurately digitized.

In praxis, in order to avoid higher frequencies of animal sounds being erroneously displayed and interpreted as lower frequencies, an anti-aliasing filter is employed in the recording system. This is a low-pass filter with a cut-off frequency below the Nyquist frequency. Frequencies higher than the Nyquist frequency are thus attenuated, so that the effect of aliasing is diminished.

An example of aliasing is given in Fig. 4.19. Spectrograms of the same killer whale (*Orcinus orca*) call are shown sampled at 96 kHz and at 32 kHz. Without an anti-aliasing filter, energy is mirror-inverted or reflected about the Nyquist frequency of 16 kHz in the second case. Conceptually, energy is folded down about the Nyquist frequency by as much as it was above the Nyquist frequency.

#### 4.3.3.3 Bit Depth

When a digitizer samples a sound wave (or the voltage at the end of a microphone), it stores the pressure measures with a limited accuracy. Bit depth is the number of bits of information in each sample. The more bits, the greater the resolution of that measure (i.e., the more accurate the pressure measure). Inexpensive sound digitizers

use 12 bits per sample. Commercially available CDs store each sample with 16 bits of storage, which allows greater accuracy in records of pressure. Blue-ray discs typically use 24 bits per sample. The more bits per sample, the larger the sound file to be stored, but the larger the dynamic range (ratio of loudest to quietest) of sounds that can be captured.

#### 4.3.3.4 Audio Coding

Audio coding is used to compress large audio files to reduce storage needs. A common format is MP3, which can achieve 75–95% file reduction compared to the original time series stored on a CD or computer hard drive. Most audio coding algorithms aim to reduce the file size while retaining reasonable quality for human listeners. The MP3 compression algorithm is based on perceptual coding, optimized for human perception, ignoring features of sound that are beyond normal human auditory capabilities. Playing MP3 files back to animals might result in quite different perception compared to the playback of the original time series. Unfortunately, this is very often ignored in animal bioacoustic experiments. Lossless compression does exist (e.g., Free Lossless Audio Codec, FLAC; see Chap. 2 on recording equipment). For animal bioacoustics research, it is best to use lossless compression or none at all.

#### 4.3.3.5 FFT Window Size (NFFT)

During Fourier analysis of a digitized sound recording, a fixed number of samples of the original time series is read and the FFT is computed on this window of samples. The number of samples is a parameter passed to the FFT algorithm and is typically represented by the variable NFFT. If NFFT samples are read from the original time series, then the Fourier transform will produce amplitude and phase measures at NFFT frequencies. However, the FFT algorithm produces a two-sided spectrum that is symmetrical about 0 Hz and contains NFFT/2 positive frequencies and NFFT/2–1 negative frequencies. To compute the power spectrum, after FFT, the amplitudes of all frequencies (positive and negative) are squared and summed. In the usual case of a time series consisting of real (i.e., not complex) numbers, the same result is obtained by doubling the squared amplitudes of the positive frequencies and discarding the negative frequencies. This means that NFFT samples in the time domain yield NFFT/2 measures in the frequency domain. The FFT values, and therefore the power spectrum calculated from them, are output at a frequency spacing:

$$\Delta f = \frac{f_s}{\mathrm{NFFT}}$$

For example, if a sound recording was sampled at 44.1 kHz and the FFT was computed over NFFT = 1024 samples, then the frequency spacing would be 43.07 Hz and the power spectrum would contain 512 frequencies: 43.07 Hz, 86.14 Hz,. . ., 22,050 Hz. A different way of looking at this is that the FFT produces spectrum levels in frequency bands of constant bandwidth. And the center frequencies in this example are 43.07 Hz, 86.14 Hz,. . ., 22,050 Hz. If there were two tones at 30 Hz and 50 Hz, then the combination of recording settings ($f_s$ = 44.1 kHz) and analysis settings (NFFT = 1024) would be unable to separate these tones. Their power would be added and reported as the single level in the frequency band centered on 43.07 Hz. To separate these two tones, a frequency spacing of no more than 20 Hz is required. This is achieved by increasing NFFT. To yield a 1-Hz frequency spacing, 1 s of recording needs to be read into the FFT; i.e., NFFT = $f_s \times 1$ s.

As the NFFT increases, the frequency spacing decreases, but at the cost of the temporal resolution. This is because an increase in NFFT means that more samples from the original time series are read in order to compute one spectrum. More samples implies that the time window over which the spectrum is computed increases. In the above example, with $f_s$ = 44.1 kHz, NFFT = 1024 samples correspond to a time window $\Delta t$ of 0.023 s:

$$\Delta t = \frac{\mathrm{NFFT}}{f_s} = \frac{1}{\Delta f}$$

While 44,100 samples last 1 s, 1024 samples only last 0.023 s. The spectrum is computed over

a time window of 0.023 s length. If the recording contained dolphin clicks of 100 µs duration, then the spectrum would be averaging over multiple clicks and ambient noise. To compute the spectrum of one click, a time window of 100 µs is desired and corresponds to NFFT $= f_s \times 100\ \mu s = 4$. This is a very short window. The resulting frequency spacing would be impractically coarse:

$$\Delta f = \frac{f_s}{\text{NFFT}} = \frac{44,100\ \text{Hz}}{4} = 10,000\ \text{Hz}$$

There is a trade-off between frequency spacing and time resolution in Fourier spectrum analysis. This is often referred to as the Uncertainty Principle (e.g., Beecher 1988): $\Delta f \times \Delta t = 1$. In spectrograms, using a large NFFT will result in sounds looking stretched out in time, while a small NFFT will result in sounds looking smudged in frequency. The combination of recording settings ($f_s$) and analysis settings (NFFT) should be optimized for the sounds of interest.

#### 4.3.3.6 FFT Window Function

The computation of a discrete Fourier transform over a finite window of samples produces spectral leakage, where some power appears at frequencies (called sidelobes) that are not part of the original time series but rather due to the length and shape of the window. If a window of samples is read off the time series and passed straight into the FFT, then the window is said to have rectangular shape. The rectangular window function has values of 1 over the length of the window and values of 0 outside (i.e., before and after). The window function is multiplied sample by sample with the original time series so that NFFT values of unaltered amplitude are passed to the FFT algorithm. A rectangular window produces a large number of sidelobes (Fig. 4.20).

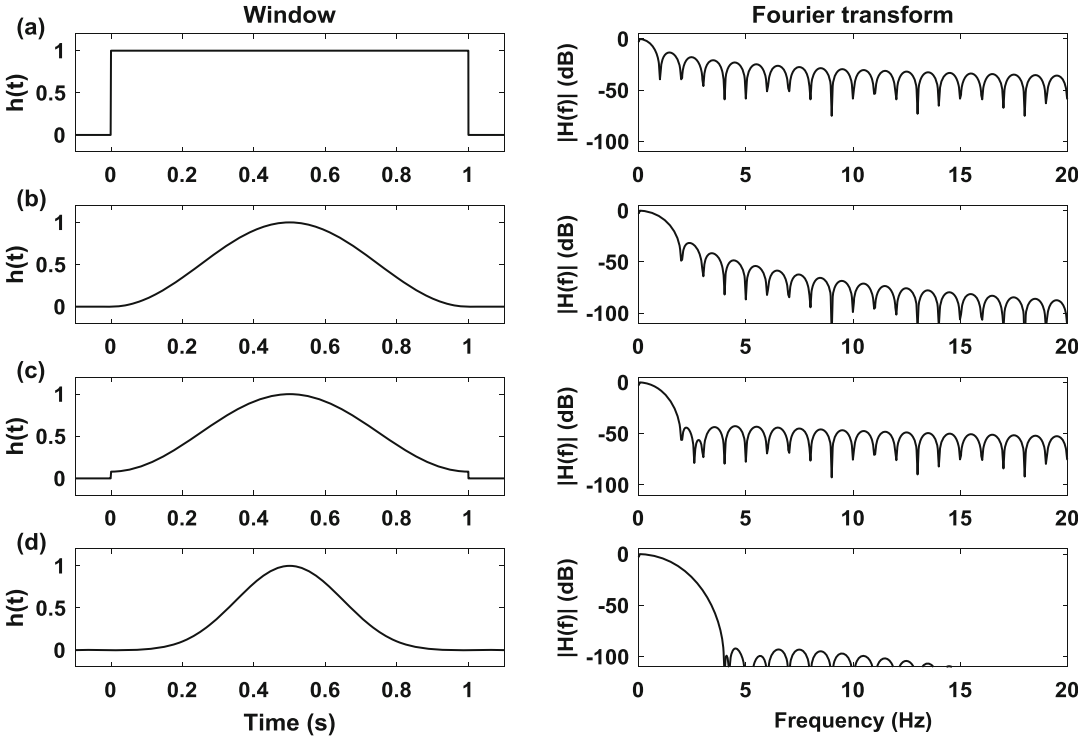



**Fig. 4.20** Comparison of some window functions (left) and their Fourier transforms (right) for (**a**) rectangular, (**b**) Hann, (**c**) Hamming, and (**d**) Blackman-Harris windows

Spectral leakage can be reduced by using non-rectangular windows such as Hann, Hamming, or Blackman-Harris windows. These have values of 1 in the center of the window, but then taper off toward the edges to values of 0. The amplitude of the original time series is thus weighted. The benefits are fewer and weaker sidelobes, which result in less spectral leakage.

The smallest difference in frequency between two tones that can be separated in the spectrum is called the frequency resolution and is determined by the width of the main lobe of the window function. There is therefore a trade-off between the reduction in sidelobes and a wider main lobe, which results in poorer frequency resolution.

In order to not miss a strong signal or strong amplitude at the edges of the window where the amplitude is weighted by values close to 0, overlapping windows are used. Rather than reading samples in adjacent windows, windows commonly have 50% overlap. A spectrogram that was computed with 50% overlapping windows will have twice the number of spectrum columns and appear to have finer time resolution. Each spectrum column still has the same $\Delta t$ as for a spectrogram without overlapping windows, but there will be twice as many spectrum columns making the spectrogram appear finer in time.

Zeros can be appended to each signal block (after windowing) to increase NFFT and therefore reduce the frequency spacing $\Delta f$. This so-called zero-padding produces a smoother spectrum but does not improve the frequency resolution, which is still determined by the shape of the window and the duration of the signal to which the window was applied.

### 4.3.4 Power Spectral Density Percentiles and Probability Density

When recording soundscapes on land or under water, sounds fade in and out, from a diversity of sources and locations. A soundscape is dynamic, changing on short to long time scales (see Chap. 7). The variability in sound levels can be expressed as power spectral density (PSD) percentiles. The $n^{th}$ percentile gives the level that is exceeded $n$% of the time (note: in engineering, the definition is commonly reversed). The $50^{th}$ percentile corresponds to the median level. An example from the ocean off southern Australia is shown in Fig. 4.21. The median ambient noise level is represented by the thin black line and goes from about 90 dB re 1 $\mu Pa^2$/Hz at 20 Hz to 60 dB re 1 $\mu Pa^2$/Hz at 30 kHz. The lowest thin gray line corresponds to the $99^{th}$ percentile. It gets quieter than this only 1% of the time. Levels at low frequencies (20–50 Hz) never drop below

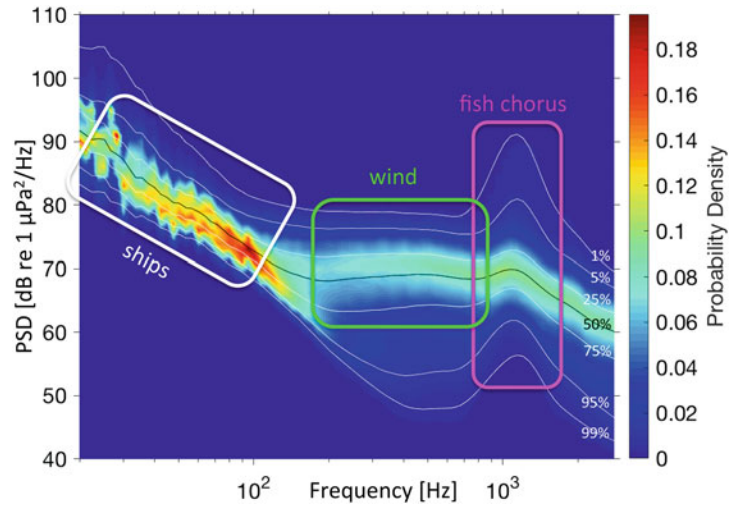



**Fig. 4.21** Percentiles of ambient noise power spectral densities measured off southern Australia over a year. Lines from top to bottom correspond to the following percentiles: 1, 5, 25, 50 (black), 75, 95, and 99

75 dB re 1 μPa$^2$/Hz because of the persistent noise from distant shipping.

These plots not only give the statistical level distribution over time, but can also identify the dominant sources in a soundscape based on the shapes of the percentile curves. The hump from 100 Hz to lower frequencies is characteristic of distant shipping. The more leveled curves at mid-frequencies (200–800 Hz) are characteristic of wind noise recorded under water. The median level of about 68 dB re 1 μPa$^2$/Hz corresponds to a Sea State of 4. The hump at 1.2 kHz is characteristic of chorusing fishes. While there are likely other sounds in this soundscape at certain times (e.g., nearby boats or marine mammals), they do not occur often enough or at a high enough level, to stand out in PSD percentile plots.

Probability density of PSD identifies the most common levels. In Fig. 4.21, at 100 Hz, the most common (probable) level was 75 dB re 1 μPa$^2$/Hz. This was equal to the median level at this frequency. The red colors indicate that the median levels were also the most probable levels. At mid-to-high frequencies, the levels were more evenly distributed (i.e., only shades of blue and no red colors). The most probable levels are not necessarily equal to the median levels. A case where the most probable level (again from distant shipping) was below the median (due to strong pygmy blue whale, *Balaenoptera musculus brevicauda*, calling) is shown in Fig. 4.6, and a case where two different levels were equally likely (due to two seismic surveys at different ranges) is shown in Fig. 4.8, both of Erbe et al. 2016a.[3] PSD percentile and probability density plots (as well as other graphs) can be created for both terrestrial and aquatic environments with the freely available software suite by Merchant et al. 2015.

## 4.4 Localization and Tracking

There are a few simple ways to gain information about the rough location and movement of a sound source. By listening in air with two ears, we can tell the direction to the sound source and whether it remains at a fixed location or approaches or departs. From recordings made over a period of time, the closest point of approach (CPA) is often taken as the point in time when mean-square pressure (or some other acoustic quantity like particle displacement, velocity, or acceleration) peaked (Fig. 4.22).

Whether a sound source is approaching or departing can also be told from the Doppler shift. As a car or a fire engine drives past and as an airplane flies overhead, the pitch drops. In fact, as each approaches, the frequency received by a listener or a recorder is higher than the emitted frequency, and as each departs, the received frequency is lower than the emitted frequency.[4] At CPA, the received frequency equals the emitted frequency. The time of CPA can be identified in spectrograms as the point in time when the steepest slope in the decreasing frequency occurred as the sound source passed or as the point in time when the frequency had decreased half-way (Fig. 4.23). The Doppler shift $\Delta f$ can easily be quantified as

$$\Delta f = \frac{v}{c} f_0$$

where $v$ is the speed of the source relative to a fixed receiver, $c$ is the speed of sound, and $f_0$ is the frequency emitted by the source (i.e., half-way between the approaching and the departing frequencies). From a spectrogram, not only the CPA, but also the speed of the sound source can be determined.

In the example of Fig. 4.23, one of the engine harmonics dropped from 96 Hz to 64 Hz. So the emitted frequency was 80 Hz and the Doppler shift was 16 Hz. With a speed of sound in air of 343 m/s, the airplane flew at 70 m/s = 250 km/h. The interesting part of this example is that the recorder was actually resting on the riverbed, in 1 m of water, and hence in a different acoustic medium to the source. How this affects the results

[3] https://www.acoustics.asn.au/conference_proceedings/AASNZ2016/papers/p14.pdf; accessed 13 October 2020.

[4] Doppler shift animations by Dan Russell: https://www.acs.psu.edu/drussell/Demos/doppler/doppler.html; accessed 13 October 2020.

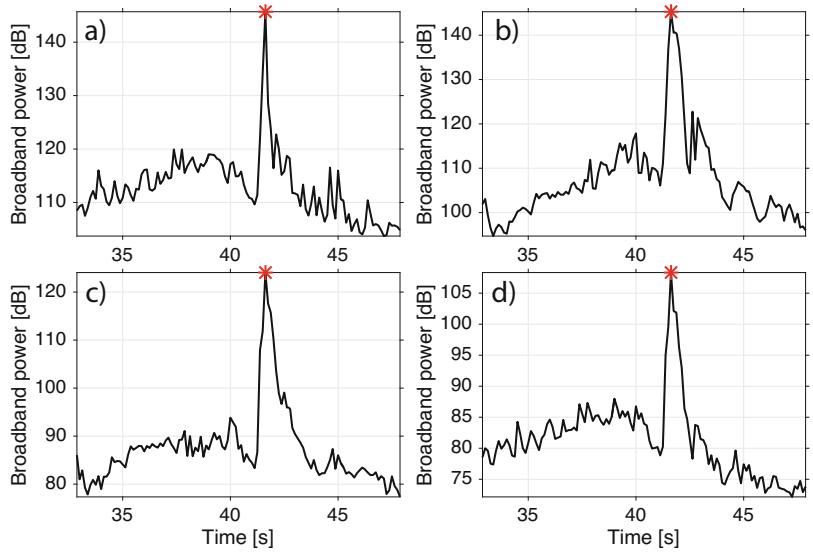



**Fig. 4.22** Graphs of (**a**) square pressure [dB re 1 μPa$^2$], (**b**) square particle displacement [dB re 1 pm$^2$], (**c**) square particle velocity [dB re 1 (nm/s)$^2$], and (**d**) square particle acceleration [dB re 1 (μm/s$^2$)$^2$] as a swimmer swims over a hydrophone. The closest point of approach is identified as the time of peak levels (i.e., at 42 s) (Erbe et al. 2017a)

depends on the depth of the hydrophone relative to the acoustic wavelength. In this particular instance, the hydrophone was a small fraction of an acoustic wavelength below the water surface and the signal reached it via the evanescent wave (see Chap. 6 on sound propagation). The evanescent wave traveled horizontally at the in-air sound speed, so it was the in-air sound speed that determined the Doppler shift. If the measurement had been carried out in deeper water with a deeper hydrophone, the signal would have been dominated by the air-to-water refracted wave, and the Doppler shift would have been determined by the in-water sound speed.

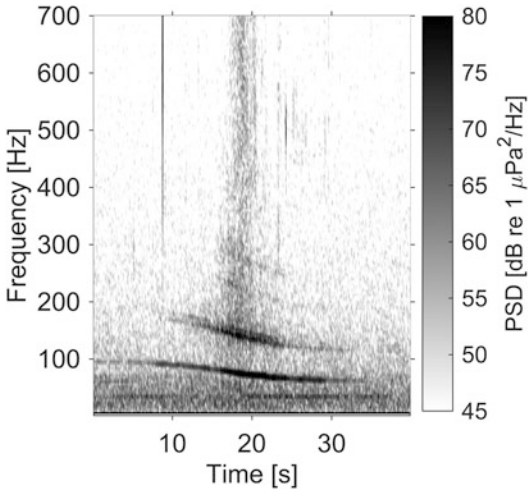



**Fig. 4.23** Spectrogram of an airplane flying over the Swan River, Perth, Australia, into Perth Airport. Recordings were made in the river, under water. The closest point of approach occurred at about 18 s, when the frequencies of the engine tone and its overtones dropped fastest (Erbe et al. 2018)

To accurately locate a sound source in space, signals from multiple simultaneous acoustic receivers need to be analyzed. These receivers are placed in specific configurations, known as arrays. Methods of localization are dependent on the configuration of the receiver array, the acoustic environment, spectral characteristics of the sound, and behavior of the sound source. There are three broad classes of these methods: time difference of arrival, beamforming, and parametric array processing methods. The following sections provide a condensed overview of the three methods. For a comprehensive treatise, please refer to the following: Schmidt 1986; Van Veen and Buckley 1988; Krim and Viberg 1996; Au and Hastings 2008; Zimmer 2011; Chiariotti et al. 2019.

Tracking is a form of passive acoustic monitoring (PAM), where an estimation of the behavior of an active sound source is maintained over time. Passive acoustic tracking has many demonstrated applications in the underwater and terrestrial domains.

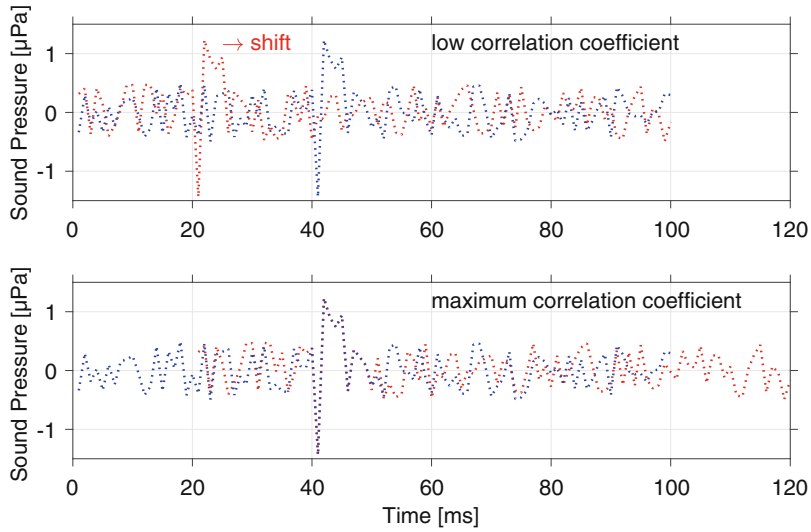



**Fig. 4.24** Determining TDOA by cross-correlation. Top: Two 100-ms time series were recorded by two spatially separated receivers. A signal of interest arrived 20 ms into the recording at receiver 1 (red) and 40 ms into the recording at receiver 2 (blue). The dot product (i.e., correlation coefficient) is low. Bottom: The red time series is shifted sample by sample against the blue time series and the dot product computed over the overlapping samples. When the signals line up, the correlation coefficient is maximum. In this example, the TDOA was 20 ms

### 4.4.1 Time Difference of Arrival

Localization by Time Difference Of Arrival (TDOA) is a two-step process. The first step is to measure the difference in time between the arrivals of the same sound at any pair of acoustic receivers. The second step is to apply appropriate geometrical calculations to locate the sound source. TDOA methods work best for signals that contain a wide range of frequencies (i.e., have a wide bandwidth), which includes short pulses, FM sweeps, and noise-like signals.

#### 4.4.1.1 Generalized Cross-Correlation

TDOAs are commonly determined by cross-correlation. The time series of recorded sound pressure by two spatially separated receivers are cross-correlated as a sliding dot product. This means that each sample from receiver 1 is multiplied with a corresponding sample from receiver 2, and the products are summed over the full length of the overlapping time series. This yields the first cross-correlation coefficient. Next, the time series from receiver 1 (red in Fig. 4.24) is shifted by 1 sample against the time series from receiver 2 (blue), and the dot product is computed again (over the overlapping samples), yielding the second cross-correlation coefficient. By sliding the two time series against each other (sample by sample) and computing the dot product, a time series of cross-correlation coefficients forms. A peak in cross-correlation occurs when the time series have been shifted such that the signal recorded by receiver 1 lines up with the signal recorded by receiver 2. The number of samples by which the time series were shifted, divided by the sampling frequency of the two receivers, is the TDOA.

Generalized cross-correlation is a common way of determining TDOA. It is suitable for localization in air and water in environments with high noise and reverberation and can be computed in either the time or frequency domains (Padois 2018).

#### 4.4.1.2 TDOA Hyperbolas

TDOAs are always computed between two receivers (from a pair of receivers). Figure 4.25 sketches the arrangement of an animal A (at point *A*) and two receivers ($R_1$ and $R_2$) in space. The

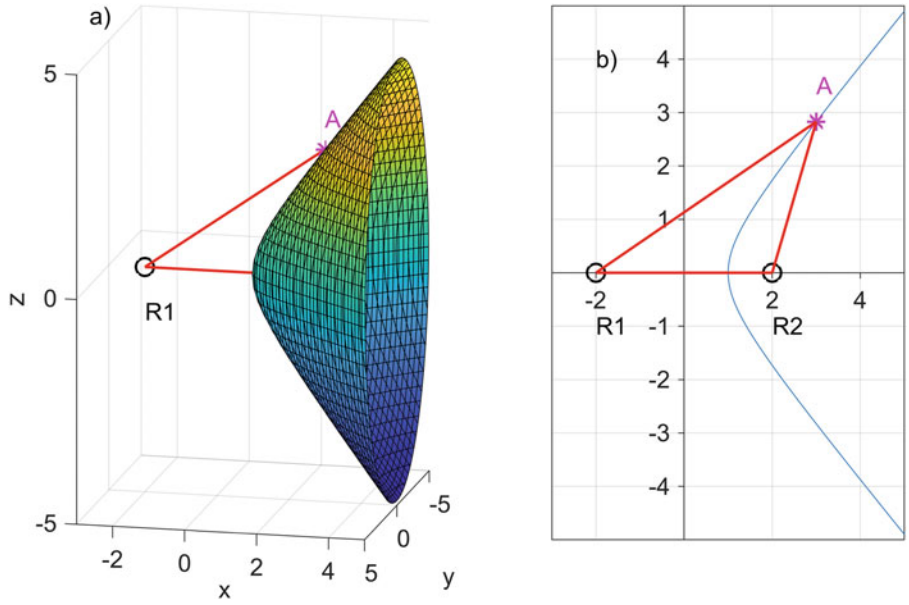



**Fig. 4.25** Graphs of localization hyperbolas with two receivers; (**a**) 3D hyperboloid and (**b**) 2D hyperbola (i.e., cross-section) in the x-z plane. A marks the animal's position; $R_1$ and $R_2$ mark the receiver positions. $R_2$ is hidden inside the hyperboloid in the 3D image

distances $A$-$R_1$ (mathematically noted as a line connecting points $A$ and $R_1$ and then taking the magnitude of it: $|\overline{A\,R_1}|$), $A$-$R_2$, and $R_1$-$R_2$ are shown as red lines. If A produces a sound that is recorded by both $R_1$ and $R_2$, then the arrival time at point $R_1$ is equal to the distance $A$-$R_1$, divided by the speed of sound $c$, and the arrival time at $R_2$ is equal to the distance $A$-$R_2$, divided by the speed of sound $c$. The TDOA is simply the difference between the two arrival times:

$$TDOA = \frac{|\overline{A\,R_1}| - |\overline{A\,R_2}|}{c}$$

It turns out mathematically that the animal can be anywhere on the hyperboloid and the TDOA will be the same. In other words, the TDOA defines a surface (in the shape of a hyperboloid) on which the animal may be located. With two receivers in the free-field, the animal's position cannot be specified further. If there are boundaries near the animal and/or receivers (e.g., if a bird is tracked with receivers on the ground), then the possible location of the animal can be easily limited (i.e., the bird cannot fly underground, eliminating half of the space). Reflections off boundaries can also be used to refine the location estimate. Finally, if one deploys more than two receivers, TDOAs can be computed between all possible pairs of receivers, yielding multiple hyperboloids that will intersect at the location of the animal.

#### 4.4.1.3 TDOA Localization in 2 Dimensions

Localization in 2D space is, of course, simpler than in 3D, though it might seem a little contrived. In Fig. 4.26, the airport arrival flight path goes straight over a home. TDOA is used to locate (and perhaps track) each airplane. Two receivers on the ground will yield the upper half of the hyperbola in Fig. 4.25b as possible airplane locations. We know the airplane cannot be underground, but in terms of its altitude and range, two receivers are unable to resolve these. A third receiver in line with $R_1$ and $R_2$ is needed. With three receivers in a line array, three TDOAs can be computed and three hyperbolas can be drawn. Any two of these hyperbolas will intersect at two points: one above and one below the x-axis (i.e., above and below ground). Knowing that the

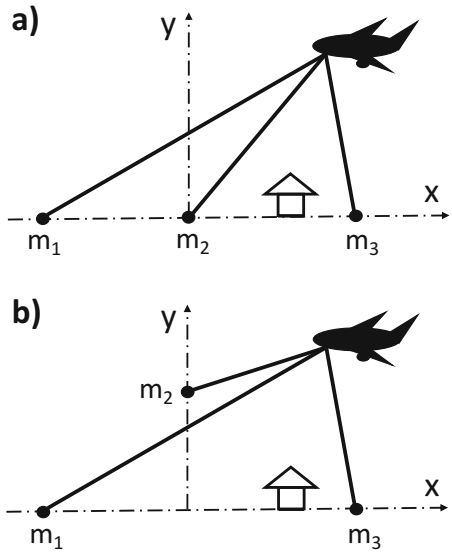



**Fig. 4.26** Sketches of a three-microphone line array (**a**) and a triangular array (**b**)

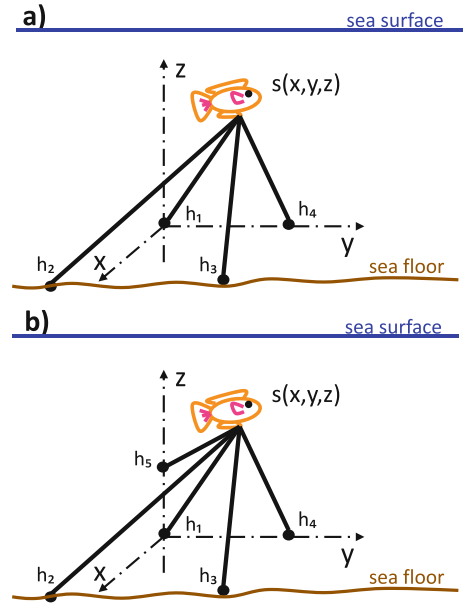



**Fig. 4.27** Sketches of seafloor-mounted arrays with 4 (**a**) and 5 (**b**) hydrophones

airplane is above ground allows its position to be uniquely determined. If there were no boundary (i.e., ground in this case), an up-down ambiguity would remain; the plane could be at either of the two intersection points. Using more than three receivers in a line array (and thus adding more TDOAs and hyperbolas) will not improve the localization capability as all hyperbolas will intersect in the same two points: one above and one below the array. The up-down ambiguity can be resolved by using a 2D rather than 1D (i.e., line) arrangement. If one microphone is moved away from the line (as in Fig. 4.26b), the TDOA hyperbolas will intersect in just one point: the exact location of the airplane.

#### 4.4.1.4 TDOA Localization in 3 Dimensions

The more common problem is to localize sound sources in 3-d space; i.e., when the sound source and the receivers are not in the same plane. Here, a line array of at least three receivers will result in hyperboloids that intersect in a circle. No matter how many receivers are in the line array, all TDOA hyperboloids will intersect in the same circle. There is up-down and left-right, in fact, circular ambiguity about the line of receivers. This is a common situation with line arrays towed behind a ship in search of marine fauna.

In order to improve localization, a fourth receiver is needed that is not in line with the others. With four receivers, three hyperboloids can be computed that will intersect in two points: one above the plane of receivers and one below, yielding another up-down ambiguity. If the receiver sits on the ground or seafloor, then one of the points can be eliminated and the sound source uniquely localized. Otherwise, a fifth hydrophone is needed that is not in the same plane as the other four, allowing general localization in 3D space (Fig. 4.27).

The dimensions of an acoustic array used for TDOA localization are determined by the expected distance to the sound source and the likely uncertainty in the TDOA measurements, which is inversely proportional to the bandwidth of the sounds being correlated. A rough estimate of the TDOA uncertainty, $\delta_t$ (s), is $\delta_t \approx 1/BW$ where $BW$ is the signal bandwidth (Hz). The corresponding uncertainty in the difference in distances from the two hydrophones to the source is then $\delta_d = c\delta_t$ where $c$ is the sound speed (m/s).

When a sound source is far away from an array of receivers, the TDOAs can still be used to determine the direction of the sound source but any estimate of its distance will become inaccurate.

### 4.4.2 Beamforming

TDOA methods give poor results for sources that emit narrow-bandwidth signals such as continuous tones (e.g., some sub-species of blue whale) and can also be confounded in situations where there are many sources of similar signals in different directions from the array (e.g., a fish chorus). However, a properly designed array can be used to determine the direction of narrowband sources and can also determine the directional distribution of sound produced by multiple, simultaneously emitting sources using a processing method called beamforming. If two or more spatially separated arrays can be deployed, then the directional information they produce can be combined to obtain a spatial localization of the source. Alternatively, if the source is known to be stationary, or moving sufficiently slowly, localization can be achieved by moving a single array, for example by towing it behind a ship.

For the convenient, and hence commonly used case of an array consisting of a line of equally spaced hydrophones, beamforming requires the hydrophone spacing to be less than half the acoustic wavelength of the sound being emitted by the source. Also, the accuracy of the bearing estimates improves as the length of the array increases. These two factors combined mean that a useful array for beamforming is likely to require at least eight hydrophones, and even that would give only modest bearing accuracy. Consequently, 16-element or even 24-element arrays are commonly deployed in practice. A straight-line array used for beamforming suffers from the same ambiguity as a TDOA array in which all the hydrophones are in a straight line. As in the TDOA case, this ambiguity can be countered by offsetting some of the hydrophones from the straight line, however beamforming requires the relative positions of all the hydrophones to be accurately known, so this is not always easy to achieve in practice.

Beamforming itself is relatively simple conceptually, but there are many subtleties (for details, see Van Veen and Buckley 1988; Krim and Viberg 1996). As for TDOA methods, the starting point is that when sound from a distant source arrives at an array of hydrophones, it will arrive at each hydrophone at a slightly different time, with the time differences depending on the direction of the sound source. The simplest type of beamformer is the delay and sum beamformer in which the array is "steered" in a particular direction by calculating the arrival time differences corresponding to that direction, delaying the received signals by amounts that cancel out those time differences, and then adding them together. This has the effect of reinforcing signals coming from the desired direction, while signals from other directions tend to cancel out. This isn't a perfect process and the array will still give some output for signals coming from other directions. The relative sensitivity of the beamformer output to signals coming from different directions can be calculated and gives the beam pattern of the array. The beam pattern of a line array depends on the steering direction, with the narrowest beams occurring when the array is steered at right-angles to the axis of the array (broadside), and the broadest beams when steered in the axial direction (end-fire). There are a number of other beamforming algorithms that can give improved performance in particular circumstances; see the above references for details.

### 4.4.3 Parametric Array Processing

The array requirements for parametric array processing methods are similar to those for beamforming, but these methods attempt to circumvent the direct dependence of the angular accuracy on the length of the array (in acoustic wavelengths) that is inherent to beamforming. A summary of these methods can be found in Krim and Viberg (1996). One of the earliest and best

known parametric methods is the multiple signal classification (MUSIC) algorithm proposed by Schmidt 1986. These methods can give more accurate localization than beamforming in situations where there is a high signal-to-noise ratio and a limited number of sources, however they are significantly more complicated to implement and more time-consuming to compute. They also rely on more assumptions and are more sensitive to errors in hydrophone positions than beamforming.

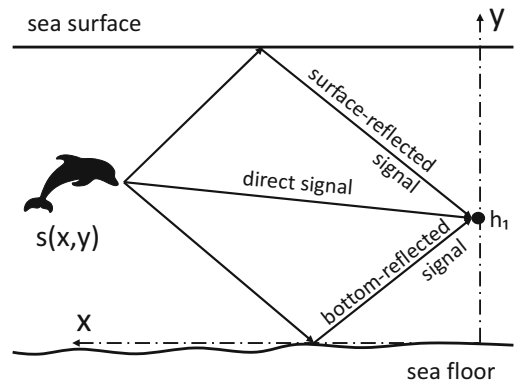



**Fig. 4.28** Sketch of localization in shallow water using a single hydrophone (Cato 1998)

### 4.4.4 Examples of Sound Localization in Air and Water

Passive acoustic localization in air poses logistical challenges with sound attenuating more rapidly in air than in water. This is an issue when localizing sound sources in open environments, as suitable recordings can only be collected if the microphone array is positioned closely around the source with localization error increasing with distance.

Sound source localization in the terrestrial domain is generally undertaken using one of three methods. Firstly, TDOA is perhaps most commonly applied to wildlife monitoring, including birds (McGregor et al. 1997) and bats (e.g., Surlykke et al. 2009; Koblitz 2018). Secondly, beamforming is more often utilized in environmental noise measurement and management (e.g., Huang et al. 2012; Prime et al. 2014; Amaral et al. 2018). Thirdly, the perhaps less common MUSIC approach has been utilized in bird monitoring and localization in noisy environments (Chen et al. 2006).

Under water, both fixed and towed hydrophone arrays are common. TDOA is the most common approach in the case of localizing cetaceans (Watkins and Schevill 1972; Janik et al. 2000) and fishes (Parsons et al. 2009; Putland et al. 2018). Under specific conditions, one or two hydrophones may suffice to localize a sound source by TDOA.

Multi-path propagation in shallow water may allow localization with just one hydrophone. TDOAs are computed between the surface-reflected, seafloor-reflected, and direct sound propagation paths yielding both range and depth of the animal (Fig. 4.28), while not being able to resolve circular symmetry (Cato 1998; Mouy et al. 2012).

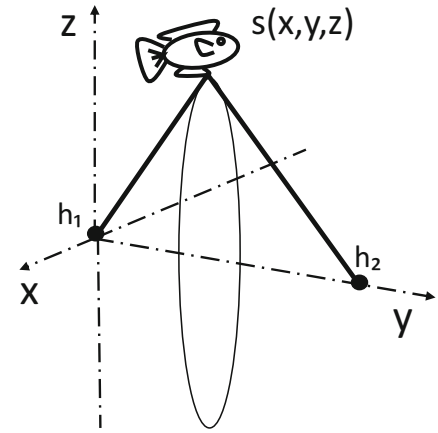



**Fig. 4.29** Sketch of two hydrophones localizing a fish in 3D space with circular ambiguity using TDOA and intensity differences (Cato 1998)

Using TDOAs in addition to differences in received intensity (when the source is located much closer to one of two receivers) may allow localization in free space to a circle between the two receivers and perpendicular to the line of two receivers (Cato 1998), see Fig. 4.29.

Beamforming is an established method for localizing soniferous marine animals (Miller and Tyack 1998) and anthropogenic sound sources

such as vessels (Zhu et al. 2018). A MUSIC approach to localization also has applications in the underwater domain, having previously been used for recovering acoustically-tagged artifacts by autonomous underwater vehicles (AUVs) (Vivek and Vadakkepat 2015).

Finally, target motion analysis involves marking the bearing to a sound source (from directional sensors or a narrow-aperture array) successively over time. If the animal calls frequently and moves slowly compared to the observation platform, successive bearings will intersect at the animal location (e.g., Norris et al. 2017).

### 4.4.5 Passive Acoustic Tracking

Passive acoustic tracking is the sequential localization of an acoustic source, useful for monitoring its behavior. Such behavior includes kinetic elements (e.g., swim path and speed) and acoustic elements (such as vocalization rate and type). In praxis, the process is a bit more complicated than just connecting TDOA locations over time. Animals will be arriving and departing; there may be more than one animal vocalizing; any one animal will have quiet times between vocalizations. So, TDOA locations need to be joined into tracks; tracks need to be continued; old tracks need to be terminated; new tracks need to be initiated; tracks may need to be merged or split. Different algorithms have been developed to aid this process, with Kalman filtering being common (Zimmer 2011; Zarchan and Musoff 2013).

While radio telemetry has historically been the primary approach to terrestrial animal tracking, passive acoustic telemetry has grown in popularity as more animals can be monitored non-invasively (e.g., McGregor et al. 1997; Matsuo et al. 2014). Passive acoustic tracking in water is a well-established method of monitoring the behavior of aquatic fauna, including their responses to environmental and anthropogenic stimuli (e.g., Thode 2005; Stanistreet et al. 2013). Both towed and moored arrays are used, with towed arrays providing greater spatial coverage in the form of line-transect surveys.

## 4.5 Symbols and Abbreviations (Table 4.10)

**Table 4.10** Most common quantities and abbreviations in this chapter

| Quantity | Abbreviation | Symbol | Unit |
|---|---|---|---|
| Frequency | | $f$ | Hz |
| Sampling frequency | | $f_s$ | Hz |
| Wavelength | | $\lambda$ | m |
| Speed of sound | | $c$ | m/s |
| Particle velocity | | $u$ | m/s |
| Period of oscillation | | $\tau$ | s |
| Time variable | | $t$ | s |
| Sound pressure | | $p(t)$ | Pa |
| Peak sound pressure | | $p_{pk}$ | Pa |
| Peak-to-peak sound pressure | | $p_{pk\text{-}pk}$ | Pa |
| Root-mean-square sound pressure | | $p_{rms}$ | Pa |
| Sound pressure level | SPL | $L_p$ | dB re 1 μPa or 20 μPa |
| Peak sound pressure level | SPLpk | $L_{p,pk}$ | dB re 1 μPa or 20 μPa |
| Radiated noise level | RNL | $L_{RN}$ | dB re 1 μPa m or 20 μPa m |
| Sound exposure level | SEL | $L_{E,p}$ | dB re 1 $\mu Pa^2 s$ or 400 $\mu Pa^2 s$ |
| Source level | SL | $L_S$ | dB re 1 or 20 μPa m |
| Number of Fourier components | NFFT | | |
| Power spectral density level | PSD | $L_{p,f}$ | dB re 1 $\mu Pa^2/Hz$ or 400 $\mu Pa^2/Hz$ |
| Time difference of arrival | TDOA | | s |

## 4.6 Summary

This chapter presented an introduction to acoustics and explained the basic quantities and concepts relevant to terrestrial and aquatic animal bioacoustics. Specific terminology that was introduced includes sound pressure, sound exposure, particle velocity, sound speed, longitudinal and transverse waves, frequency modulation, amplitude modulation, decibel, source level, near-field, far-field, frequency weighting, power spectral density, and one-third octave band level, amongst others. The chapter further introduced basic signal sampling and processing concepts such as sampling frequency, Nyquist frequency, aliasing, windowing, and Fourier transform. The chapter concluded with an introductory treatise of sound localization and tracking, including time difference of arrival and beamforming.

## References


Acton WI (1974) The effects of industrial airborne ultrasound on humans. Ultrasonics 12(3):124–128. https://doi.org/10.1016/0041-624X(74)90069-9

Amaral FR, Serrano Rico JC, Medeiros MAF (2018) Design of microphone phased arrays for acoustic beamforming. J Braz Soc Mech Sci Eng 40(7):354. https://doi.org/10.1007/s40430-018-1275-5

American National Standards Institute (2013) Acoustical Terminology (ANSI/ASA S1.1-2013). Acoustical Society of America, Melville, NY, USA

Au WWL, Hastings M (2008) Principles of marine bioacoustics. Springer Verlag, New York

Baumann-Pickering S, McDonald MA, Simonis AE, Solsona Berga A, Merkens KPB, Oleson EM, Roch MA, Wiggins SM, Rankin S, Yack TM, Hildebrand JA (2013) Species-specific beaked whale echolocation signals. J Acoust Soc Am 134(3):2293–2301. https://doi.org/10.1121/1.4817832

Beecher MD (1988) Spectrographic analysis of animal vocalizations: Implications of the "Uncertainty Principle". Bioacoustics 1(2–3):187–208. https://doi.org/10.1080/09524622.1988.9753091

Caldwell MC, Caldwell DK (1965) Individualized whistle contours in bottlenosed dolphins (*Tursiops truncatus*). Nature 207(4995):434–435. https://doi.org/10.1038/207434a0

Cato DH (1998) Simple methods of estimating source levels and locations of marine animal sounds. J Acoust Soc Am 104(3):1667–1678. https://doi.org/10.1121/1.424379

Chen C-E, Ali AM, Wang H (2006) Design and testing of robust acoustic arrays for localization and enhancement of several bird sources. Proceedings of the 5th International Conference on Information Processing in Sensor Networks:268–275

Chiariotti P, Martarelli M, Castellini P (2019) Acoustic beamforming for noise source localization – Reviews, methodology and applications. Mech Syst Signal Process 120:422–448. https://doi.org/10.1016/j.ymssp.2018.09.019

Dziak RP, Bohnenstiehl DR, Matsumoto H, Fox CG, Smith DK, Tolstoy M, Lau T-K, Haxel JH, Fowler MJ (2004) P- and T-wave detection thresholds, Pn velocity estimate, and detection of lower mantle and core P-waves on ocean sound-channel hydrophones at the Mid-Atlantic Ridge. Bull Seismol Soc Am 94(2):665–677. https://doi.org/10.1785/0120030156

Erbe C (2009) Underwater noise from pile driving in Moreton Bay, Qld. Acoust Aust 37(3):87–92

Erbe C (2013) Underwater noise of small personal watercraft (jet skis). J Acoust Soc Am 133(4):EL326–EL330. https://doi.org/10.1121/1.4795220

Erbe C, King AR (2009) Modelling cumulative sound exposure around marine seismic surveys. J Acoust Soc Am 125(4):2443–2451. https://doi.org/10.1121/1.3089588

Erbe C, McCauley R, Gavrilov A, Madhusudhana S, Verma A (2016a). The underwater soundscape around Australia. Proceedings of Acoustics 2016, 9–11 November 2016, Brisbane, Australia.

Erbe C, Parsons M, Duncan AJ, Allen K (2016b) Underwater acoustic signatures of recreational swimmers, divers, surfers and kayakers. Acoust Aust 44(2):333–341. https://doi.org/10.1007/s40857-016-0062-7

Erbe C, Parsons M, Duncan AJ, Lucke K, Gavrilov A, Allen K (2017a) Underwater particle motion (acceleration, velocity and displacement) from recreational swimmers, divers, surfers and kayakers. Acoust Aust 45:293–299. https://doi.org/10.1007/s40857-017-0107-6

Erbe C, Parsons M, Duncan AJ, Osterrieder S, Allen K (2017b) Aerial and underwater sound of unmanned aerial vehicles (UAV, drones). J Unmanned Veh Syst 5(3):92–101. https://doi.org/10.1139/juvs-2016-0018

Erbe C, Williams R, Parsons M, Parsons SK, Hendrawan IG, Dewantama IMI (2018) Underwater noise from airplanes: An overlooked source of ocean noise. Mar Pollut Bull 137:656–661. https://doi.org/10.1016/j.marpolbul.2018.10.064

Finneran J, Schlundt C (2011) Subjective loudness level measurements and equal loudness contours in a bottlenose dolphin (*Tursiops truncatus*). J Acoust Soc Am 130(5):3124–3136. https://doi.org/10.1121/1.3641449

Fletcher H, Munson WA (1933) Loudness, its definition, measurement and calculation. J Acoust Soc Am 5(2):82–108. https://doi.org/10.1121/1.1915637

Holland RA, Waters DA, Rayner JMV (2004) Echolocation signal structure in the Megachiropteran bat

*Rousettus aegyptiacus* Geoffroy 1810. J Exp Biol 207(25):4361. https://doi.org/10.1242/jeb.01288

Houser DS, Yost W, Burkard R, Finneran JJ, Reichmuth C, Mulsow J (2017) A review of the history, development and application of auditory weighting functions in humans and marine mammals. J Acoust Soc Am 141(3):1371–1413. https://doi.org/10.1121/1.4976086

Huang X, Bai L, Vinogradov I, Peers E (2012) Adaptive beamforming for array signal processing in aeroacoustic measurements. J Acoust Soc Am 131(3): 2152–2161. https://doi.org/10.1121/1.3682041

International Electrotechnical Commission (2013) Electroacoustics - Sound level meters - Part 1: Specifications (IEC 61672-1 Ed. 2.0 b:2013). New York.

International Organization for Standardization (2003) Acoustics—Normal equal-loudness-level contours (ISO 226:2003).

International Organization for Standardization (2007) Acoustics — Definitions of basic quantities and terms (ISO/TR 25417). Geneva, Switzerland.

International Organization for Standardization (2017) Underwater acoustics—Terminology (ISO 18405). Geneva, Switzerland.

Janik VM, Van Parijs SM, Thompson PM (2000) A two-dimensional acoustic localization system for marine mammals (Note). Mar Mamm Sci 16(2):437–447. https://doi.org/10.1111/j.1748-7692.2000.tb00935.x

Kamminga C, Beitsma GR (1990) Investigations on cetacean sonar IX, remarks on dominant sonar frequencies from *Tursiops truncatus*. Aquat Mamm 16(1):14–20

Kastelein RA, Wensveen PJ, Terhune JM, de Jong CAF (2011) Near-threshold equal-loudness contours for harbour seals (*Phoca vitulina*) derived from reaction times during underwater audiometry: a preliminary study. J Acoust Soc Am 129(1):488–495. https://doi.org/10.1121/1.3518779

Koblitz JC (2018) Arrayvolution: using microphone arrays to study bats in the field. Can J Zool 96(9):933–938. https://doi.org/10.1139/cjz-2017-0187

Krim H, Viberg M (1996) Two decades of array signal processing research: the parametric approach. IEEE Signal Process Mag 13(4):67–94. https://doi.org/10.1109/79.526899

Kuehne LM, Erbe C, Ashe E, Bogaard LT, Collins MS, Williams R (2020) Above and below: Military aircraft noise in air and under water at Whidbey Island, Washington. J Mar Sci Eng 8(11):923. https://doi.org/10.3390/jmse8110923

Leighton TG (2018) Ultrasound in air—Guidelines, applications, public exposures, and claims of attacks in Cuba and China. J Acoust Soc Am 144(4): 2473–2489. https://doi.org/10.1121/1.5063351

Marley SA, Erbe C, Salgado Kent CP (2017) Underwater recordings of the whistles of bottlenose dolphins in Fremantle Inner Harbour, Western Australia. Sci Data 4(170):126. https://doi.org/10.1038/sdata.2017.126

Matsuo I, Wheeler A, Kloepper L, Gaudette J, Simmons JA (2014) Acoustic tracking of bats in clutter environments using microphone arrays. J Acoust Soc Am 135(4):2207–2207. https://doi.org/10.1121/1.4877207

McGregor PK, Dabelsteen T, Clark CW, Bower JL, Holland J (1997) Accuracy of a passive acoustic location system: empirical studies in terrestrial habitats. Ethol Ecol Evol 9(3):269–286. https://doi.org/10.1080/08927014.1997.9522887

Merchant ND, Fristrup KM, Johnson MP, Tyack PL, Witt MJ, Blondel P, Parks SE (2015) Measuring acoustic habitats. Methods Ecol Evol 6(3):257–265. https://doi.org/10.1111/2041-210X.12330

Miller PJ, Tyack PL (1998) A small towed beamforming array to identify vocalizing resident killer whales (*Orcinus orca*) concurrent with focal behavioral observations. Deep Sea Res Part II Top Stud Oceanogr 45(7):1389–1405. https://doi.org/10.1016/S0967-0645(98)00028-9

Mouy X, Hannay D, Zykov M, Martin B (2012) Tracking of Pacific walruses in the Chukchi Sea using a single hydrophone. J Acoust Soc Am 131(2):1349–1358. https://doi.org/10.1121/1.3675008

Norris TF, Dunleavy KJ, Yack TM, Ferguson EL (2017) Estimation of minke whale abundance from an acoustic line transect survey of the Mariana Islands. Mar Mamm Sci. https://doi.org/10.1111/mms.12397

Padois T (2018) Acoustic source localization based on the generalized cross-correlation and the generalized mean with few microphones. J Acoust Soc Am 143(5): EL393–EL398. https://doi.org/10.1121/1.5039416

Parrack HO (1966) Effect of Air-Borne Ultrasound on Humans. Int J Audiol 5(3):294–308. https://doi.org/10.3109/05384916609074198

Parsons MJ, McCauley RD, Mackie MC, Siwabessy P, Duncan AJ (2009) Localization of individual mulloway (*Argyrosomus japonicus*) within a spawning aggregation and their behaviour throughout a diel spawning period. ICES J Mar Sci 66(6):1007–1014. https://doi.org/10.1093/icesjms/fsp016

Prime Z, Doolan C, Zajamsek B (2014) Beamforming array optimisation and phase averaged sound source mapping on a model wind turbine. Inter-Noise and Noise-Con Congress and Conference Proceedings 249(7):1078–1086

Putland RL, Mackiewicz AG, Mensinger AF (2018) Localizing individual soniferous fish using passive acoustic monitoring. Ecol Inform 48:60–68. https://doi.org/10.1016/j.ecoinf.2018.08.004

Robinson DW, Dadson RS (1956) A re-determination of the equal-loudness relations for pure tones. Br J Appl Phys 7(5):166–181. https://doi.org/10.1088/0508-3443/7/5/302

Schmidt R (1986) Multiple emitter location and signal parameter estimation. IEEE Trans Antennas Propag 34(3):276–280. https://doi.org/10.1109/TAP.1986.1143830

Southall BL, Bowles AE, Ellison WT, Finneran JJ, Gentry RL, Greene CRJ, Kastak D, Ketten DR, Miller JH, Nachtigall PE, Richardson WJ, Thomas JA, Tyack

PL (2007) Marine mammal noise exposure criteria: Initial scientific recommendations. Aquat Mamm 33(4):411–521. https://doi.org/10.1080/09524622.2008.9753846

Southall BL, Finneran JJ, Reichmuth C, Nachtigall PE, Ketten DR, Bowles AE, Ellison WT, Nowacek DP, Tyack PL (2019) Marine mammal noise exposure criteria: Updated scientific recommendations for residual hearing effects. Aquat Mamm 45(2):125–232. https://doi.org/10.1578/AM.45.2.2019.125

Stanistreet JE, Risch D, Van Parijs SM (2013) Passive acoustic tracking of singing humpback whales (*Megaptera novaeangliae*) on a Northwest Atlantic feeding ground. PLoS One 8(4):e61263. https://doi.org/10.1371/journal.pone.0061263

Surlykke A, Pedersen SB, Jakobsen L (2009) Echolocating bats emit a highly directional sonar sound beam in the field. Proc R Soc B Biol Sci 276(1658):853–860. https://doi.org/10.1098/rspb.2008.1505

Suzuki Y, Takeshima H (2004) Equal-loudness-level contours for pure tones. J Acoust Soc Am 116(2):918–933. https://doi.org/10.1121/1.1763601

Taylor B, Thompson A (eds) (2008) The International System of Units (SI). National Institute of Standards and Technology, Gaithersburg, MD

Thode A (2005) Three-dimensional passive acoustic tracking of sperm whales (*Physeter macrocephalus*) in ray-refracting environments. J Acoust Soc Am 118(6):3575–3584. https://doi.org/10.1121/1.2049068

Tonin R (2018) A review of wind turbine-generated infrasound: Source, measurement and effect on health. Acoust Aust 46(1):69–86. https://doi.org/10.1007/s40857-017-0098-3

Tougaard J, Beedholm K (2019) Practical implementation of auditory time and frequency weighting in marine bioacoustics. Appl Acoust 145:137–143. https://doi.org/10.1016/j.apacoust.2018.09.022

Van Veen BD, Buckley KM (1988) Beamforming: a versatile approach to spatial filtering. IEEE ASSP Mag 5(2):4–24. https://doi.org/10.1109/53.665

Vivek R, Vadakkepat P (2015) Multiple signal classification (MUSIC) based underwater acoustic localization module (UALM) for AUV. Paper presented at the 2015 IEEE Underwater Technology (UT) Conference. https://doi.org/10.1109/UT.2015.7108288.

Ward R, Parnum I, Erbe C, Salgado-Kent CP (2016) Whistle characteristics of Indo-Pacific bottlenose dolphins (*Tursiops aduncus*) in the Fremantle Inner Harbour, Western Australia. Acoust Aust 44(1):159–169. https://doi.org/10.1007/s40857-015-0041-4

Watkins WA, Schevill WE (1972) Sound source location by arrival-times on a non-rigid three-dimensional hydrophone array. Deep-Sea Res Oceanogr Abstr 19(10):691–706. https://doi.org/10.1016/0011-7471(72)90061-7

Wellard R, Erbe C, Fouda L, Blewitt M (2015) Vocalisations of killer whales (*Orcinus orca*) in the Bremer Canyon, Western Australia. PLoS One 10(9):e0136535. https://doi.org/10.1371/journal.pone.0136535

Zarchan P, Musoff H (2013) Fundamentals of Kalman filtering: a practical approach, 4th edn. American Institute of Aeronautics and Astronautics, Inc., Reston, VA

Zhu C, Garcia H, Kaplan A, Schinault M, Handegard NO, Godø OR, Huang W, Ratilal P (2018) Detection, localization and classification of multiple mechanized ocean vessels over continental-shelf scale regions with passive ocean acoustic waveguide remote sensing. Remote Sens 10(11):1699

Zimmer WMX (2011) Passive acoustic monitoring of cetaceans. Cambridge University Press, Cambridge

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



# Source-Path-Receiver Model for Airborne Sounds

5

Ole Næsbye Larsen, William L. Gannon, Christine Erbe, Gianni Pavan, and Jeanette A. Thomas

## 5.1 Introduction

The source-path-receiver model (SPRM) provides a common framework for occupational health and safety management. It is used for hazard control to minimize the risk of exposing workers to hazards. Such hazards may be chemicals (e.g., spilled compounds in a pharmaceutical laboratory), material (e.g., falling bricks on a construction site), or noise.

An example SPRM for chemical hazards is shown in Fig. 5.1a. The source is a poisonous chemical, which leaks through the air inside a laboratory, and the receiver is a pharmaceutical worker. The SPRM guides the health and safety manager in minimizing the risk of exposure.[1] Ideally, the source would be eliminated, but this might not be possible if this type of chemical is required. Maybe it can be substituted by a less volatile or toxic chemical? There may be engineering controls such as installing an isolation chamber (or glove box) or exhaust hood. Engineering controls may also be applied to the path along which the chemical travels: installing ventilators, absorbing material, or mechanical barriers, or simply extending the length of the path to increase dilution. Finally, controls may be applied at the receiver: proper training for safe handling of the chemical, limiting work hours, rotating shifts, and wearing personal protective equipment (PPE). In terms of reducing the risk of exposure, the measures rank from most to least effective (termed hierarchy of control): elimination, engineering controls, procedural controls, and finally, PPE.

The SPRM applied to noise control helps break down the components of noise exposure that can be modified to reduce the risk of acoustic impacts. In the example of Fig. 5.1b, the source is a busy downtown road. Noise from the cars

Jeanette A. Thomas (deceased) contributed to this chapter while at the Department of Biological Sciences, Western Illinois University-Quad Cities, Moline, IL, USA

O. N. Larsen (✉)
Department of Biology, University of Southern Denmark, Odense M, Denmark
e-mail: onl@biology.sdu.dk

W. L. Gannon
Department of Biology, Museum of Southwestern Biology, and Graduate Studies, University of New Mexico, Albuquerque, NM, USA
e-mail: wgannon@unm.edu

C. Erbe
Centre for Marine Science and Technology, Curtin University, Perth, WA, Australia
e-mail: c.erbe@curtin.edu.au

G. Pavan
Department of Earth and Environment Sciences, University of Pavia, Pavia, Italy
e-mail: gianni.pavan@unipv.it

[1] Example SPRM for hazard control. Canadian Centre for Occupational Health and Safety, Government of Canada; https://www.ccohs.ca/oshanswers/hsprograms/hazard_control.html; accessed 4 December 2020.

© The Author(s) 2022
C. Erbe, J. A. Thomas (eds.), *Exploring Animal Behavior Through Sound: Volume 1*,
https://doi.org/10.1007/978-3-030-97540-1_5

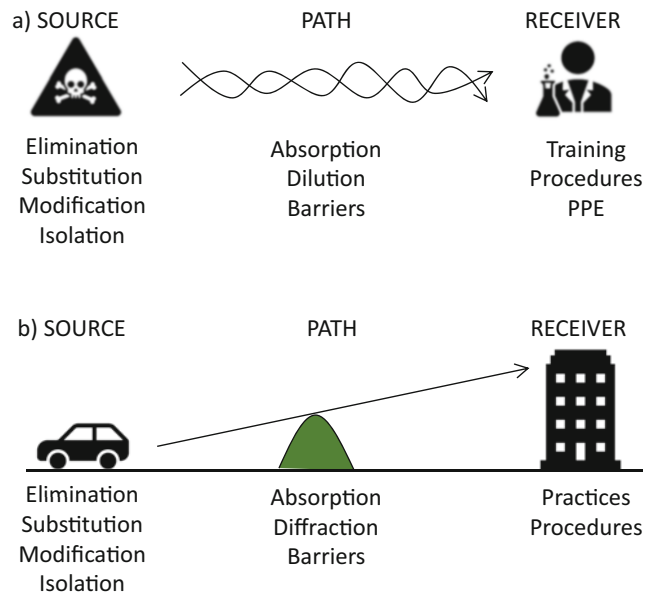



**Fig. 5.1** Examples of the source-path-receiver model for (**a**) chemical hazard control in a laboratory and (**b**) traffic noise control in a city

travels to surrounding residential buildings.[2] The source may be eliminated by relocating all traffic to an inner-city bypass and banning all traffic downtown. Maybe private car traffic can be substituted by a quieter, electric city bus service. Imposing a speed limit reduces noise. Some cities enforce noise emission standards for cars. Long-term engineering solutions may include building a tunnel, resurfacing the road with noise-absorbing material, installing noise barrier walls along the road, or erecting earth bunds. Residential buildings may have noise-reduction (double-glazed) windows and residents may set up their bedrooms at the opposite side of the building. The specific implementation of the SPRM depends on the application. For example, residents in an apartment building would not want to wear earmuffs at home, but for workers in a noisy plant, such PPE is common practice. A poster showing the steps involved in workplace noise control is shown in Fig. 5.2.

Even though the SPRM was originally developed to manage hazards at the workplace, it is much more broadly applicable to the day-to-day lives of humans—and animals. In fact, the SPRM is fundamental. Without a receiver, there is no hazard. Without a listener, there is no noise. Researchers of animal bioacoustics might want to apply the SPRM to their project in order to identify parameters of the source, path, and receiver, that might influence the results. Other chapters in this book either explicitly or implicitly apply the SPRM. Chapter 13 on the effects of noise on animals provides examples where the source is a highway, the path follows from the highway into the surrounding bush, and the receivers are birds, whose abundance might decrease closer to the source as a result of habitat degradation by noise. Chapter 11 deals with acoustic communication between animals, and so the source may be a male frog, the path may lead through a tropical rain forest, and the receivers are nearby females of the same species. Chapter 12 is about echolocation. Here, the source and the receiver are the same individual animal. A bat echolocates on a moth and the echolocation signal reflects off the moth,

[2] Example SPRM for traffic noise. Environmental Protection Department, The Government of the Hong Kong Special Administrative Region https://www.epd.gov.hk/epd/noise_education/young/eng_young_html/m3/m3.html; accessed 4 December 2020.

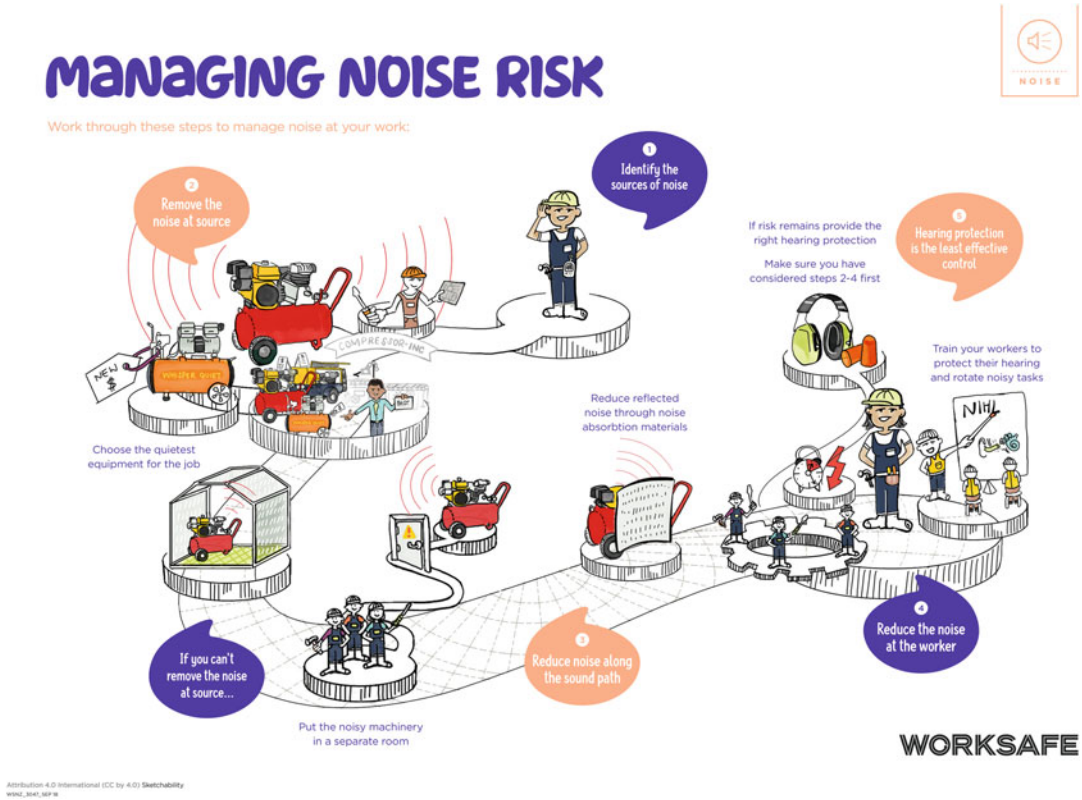



**Fig. 5.2** Poster by WorkSafe New Zealand illustrating the steps involved in noise control at the workplace. © WorkSafe, New Zealand Government, 2018; https://www.worksafe.govt.nz/dmsdocument/3987-managing-noise-risk-poster. Reproduced with permission; https://www.worksafe.govt.nz/about-us/about-this-site/copyright/. A more elaborate animation is also available (Animation of the SPRM by WorkSafe, New Zealand Government; https://youtu.be/8Cq5UR5KssA; accessed 4 December 2020.)

informing the bat how far away its prey is. The signal travels through the environment twice: from the bat to its prey and back. Chapter 10 covers audiometry, where the sources are controlled and engineered signals (often pure tones) that are played to animals over short distances or through earphones, and the receivers are individual animals whose hearing is being measured. Chapter 7 explores soundscapes on land and under water. The sources are grouped into geophony (e.g., wind, rain, and waves), biophony (i.e., animals), and anthropophony (e.g., airplanes or ships). The paths go through the air over land, under water, and through the ground. The receivers in passive acoustic monitoring of soundscapes are recorders, which collect and store acoustic data for later analysis in the laboratory. The following sections first explore the basic concepts of sound propagation in air before applying these to an example SPRM.

## 5.2 Sound Propagation in Terrestrial Environments

The environment through which a sound travels alters its acoustic features such as its spectral composition and level. The effects of the environment on bioacoustic signals were well explored in the classic works of Chappuis (1971), Marten and Marler (1977), Michelsen (1978), and Wiley and Richards (1978).

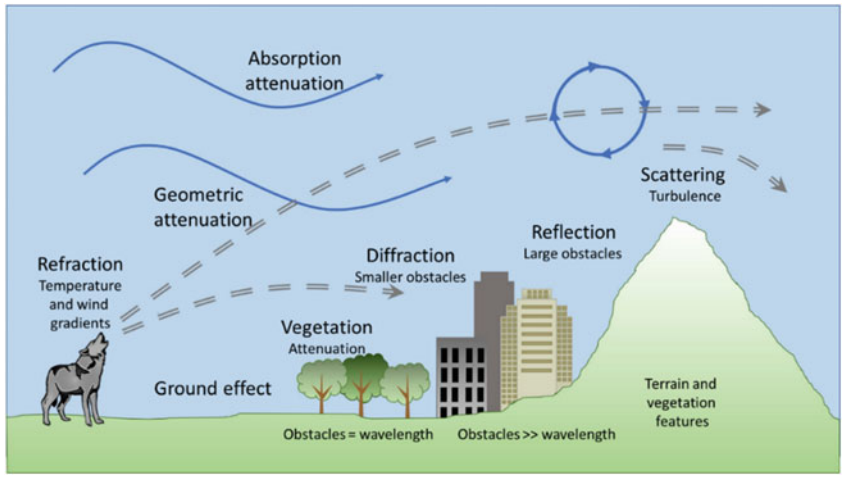



**Fig. 5.3** Diagram of some of the factors affecting sound propagation in air. Figure donated by Sara Torres Ortiz

Airborne sound propagation (often called outdoor sound propagation) is characterized by a number of phenomena. Sounds attenuate with distance from the sender due to geometrical attenuation (i.e., spreading) and absorption by the medium. High-frequency sounds (i.e., sounds having short wavelengths; see Chap. 4 on definitions of frequency and wavelength) propagate over shorter distances than low-frequency sounds (i.e., sounds having long wavelengths). Environmental and structural factors such as substrate composition; terrain profile; obstacles along the path; amount of vegetative cover; wind speed and direction; vertical gradients (i.e., increases or decreases) in wind speed, air temperature, and humidity; air turbulence; and, to a small degree, altitude (i.e., atmospheric pressure) affect sound propagation in air (Fig. 5.3). The propagation paths, along which sounds travel, are rarely straight lines, but rather bend (i.e., refract or diffract), reflect, and scatter. The same sound traveling along different propagation paths may interfere with itself constructively or destructively. The received sound is a weaker and often distorted version of the sent sound (Wahlberg and Larsen 2017).

This section explains the basic concepts of sound propagation in air and provides some insights into environmental effects on propagation. Some environmental factors (e.g., air temperature, wind speed and direction, and humidity) vary throughout the day and among seasons, and so sound propagation can be quite variable. Sound propagation models exist and can be used to predict the distance over which sounds travel, create noise maps, estimate changes to the acoustic (e.g., spectral) features of received sounds, and identify factors that could hinder or enhance animal communication (see Lohr et al. 2003; Jensen et al. 2008). Bioacousticians should consider the characteristics of sound propagation, which could explain variability in the receiver's behavioral response or the effectiveness of acoustic communication.

### 5.2.1 Ray Traces

Sound propagation is accurately described by the acoustic wave equation. This is a four-dimensional (4-d: three spatial coordinates and time) differential equation of the second order. For an "easy" derivation of the acoustic wave equation, see Larsen and Radford (2018). However, in the simplest situation of symmetric geometry (i.e., omnidirectional signal in a homogeneous medium with no reverberation), the equation can be simplified and described by one variable: the range to the source (Wahlberg and Larsen 2017). Even then, solving the wave

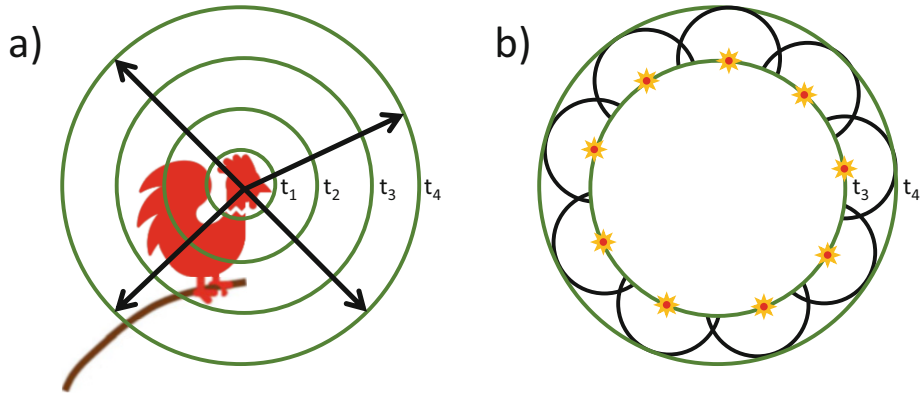



**Fig. 5.4** (**a**) Sketch of a rooster sitting on a branch. When the bird crows, sound is emitted in all directions (marked by a few example black arrows). The green concentric circles represent the wavefronts of the outgoing sound at times $t_1 - t_4$. The wave rays are perpendicular to the wavefronts and point in the direction of sound propagation. (**b**) Illustration of Huygens' principle. Each point on the wavefront at time $t_4$ can be considered itself a (secondary) source; nine example points are marked by suns. The wavefronts of the secondary sources (shown as black circle segments) superpose to yield the new primary wavefront, drawn at time $t_4$

equation under the various and variable conditions encountered in common sound propagation scenarios is quite a task. Fortunately, there are much simpler, conceptual principles of sound propagation, which can yield satisfactory results. One such concept is ray propagation or ray tracing.

Let us consider an omnidirectional source, which emits sound equally in all directions. An example is the crowing rooster in Fig. 5.4a (although it is only omnidirectional at the lower frequencies of its crow and it might not typically crow while roosting, but for the sake of science. . .; Larsen and Dabelsteen 1990). Wave rays point in the direction of sound propagation and are perpendicular to the wavefronts of the propagating sound. The wavefronts are spheres in 3D space (circles in 2D). Huygens' principle (named after Christiaan Huygens, a Dutch physicist) states that every point on a wavefront can be considered a source of a new (secondary) wave. And all of the secondary wavefronts superpose to build the next (in time) primary wavefront. The wavefront at time $t_3$ in Fig. 5.4a is also shown in Fig. 5.4b. Nine example points on this wavefront are "randomly" illustrated (as small suns). These each create their own set of concentric wavefronts, drawn at time $t_4$. The secondary waves cancel out in some places but at the farthest range from the rooster in the center, the secondary wavefronts line up to yield the new primary wavefront at time $t_4$.

As the expanding wavefront encounters features of the environment (e.g., vegetation or gradients in sound speed), its shape changes and the directions of the wave rays change. The laws of physics and principles of sound propagation can be applied to trace the propagation paths. This is called ray tracing. For an easy introduction to ray tracing, see Heller (2013). Wahlberg and Larsen (2017) suggested visualizing a ray as a "small acoustic particle travelling along a narrow beam or ray in discrete steps and bouncing-off or being refracted through surfaces." This type of sound field visualization, first introduced in 1967 (Krokstad et al. 2015), has been used extensively in linear acoustics to model phenomena in outdoor sound propagation with the computational tools now available with computers (Attenborough et al. 1995).

An example of ray tracing is shown in Fig. 5.5. The omnidirectional source is located in the lower left corner, 5 m above ground at range 0, and it emits a 10-Hz tone. The wave rays are shown and follow the sound propagation paths. Sound that is initially emitted in an upwards direction bends

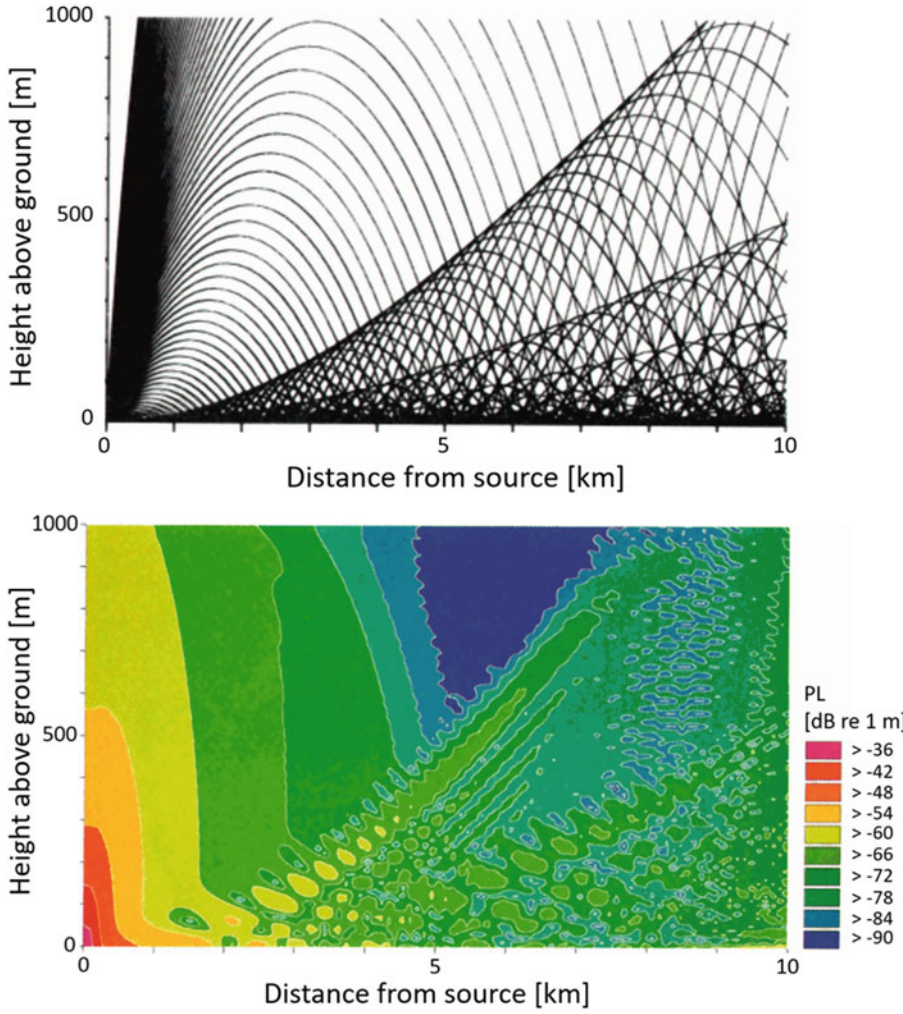



**Fig. 5.5** Top: Ray traces modeling the propagation of an airborne 10-Hz tone from a point source located 5 m off the ground (lower left corner). The model suggests that sound is bent downwards (downward refraction, typical for nighttime) where it bounces off the ground several times depending on the initial direction from the source. Note the scales: These effects occur at distances much longer than typical animal sound communication distances, which normally are up to only a few hundred meters. Bottom: Contour plot of propagation loss, PL (i.e., attenuation) of the 10-Hz sound. Modified from Attenborough et al. (1995). © Acoustical Society of America, 1995. All rights reserved

downward at a certain altitude (depending on its initial angle of emission). This is typical for nighttime sound propagation. Once rays hit the ground, they are reflected upwards again. The sound field (i.e., the received level at every location in space) is computed by summing sound pressure over all rays. Regions where rays travel close together have high received levels (little propagation loss) and regions that only a few rays enter have low received levels (high propagation loss).

For example, Ottemöller and Evers (2008) used ray tracing to describe the sound propagation of a massive vapor cloud explosion at Buncefield fuel depot near Hemel Hempstead, UK, on the morning of 11 December 2005. The storage tank overflowed and released over

300 tons of fuel. An explosion was triggered after a vapor cloud formed and spread over a very large area (80,000 $m^2$ or about 20 acres) before igniting. The explosion was huge, caused extensive damage, injured 43 people, and was detected by seismograph stations in the UK and the Netherlands. The data provided significant information on the ray trajectories of this explosion.

### 5.2.2 Geometrical Sound Spreading

Sound from an omnidirectional source in the free-field spreads out evenly in a spherical pattern (i.e., equally in all directions). The free-field is homogeneous (i.e., has no temperature or humidity gradients) and unimpeded by buildings or vegetation. At any receiver location in space, only a small proportion of the emitted sound arrives, and so the received sound is attenuated compared to the sound energy emitted at the source. The total attenuation or loss of sound energy from the source to a receiver is known as propagation loss (PL; formerly transmission loss). The sound pressure level at the source (defined as 1 m from a point source; see Chap. 4) is called the source level (SL), whereas the sound pressure level at the receiver at a distance (i.e., range $r$) from the source is called the received level (RL). The relation between these two levels is given by Eq. 5.1:

$$RL = SL - PL \tag{5.1}$$

Propagation loss in the free-field is termed spherical spreading loss, which can be computed as $PL_{sph} = 20 \log_{10}(r)$ (for derivation of this expression, see Wahlberg and Larsen 2017). It is independent of signal frequency and only depends on the geometry of the source and sound field. So, Eq. 5.1 may be reformulated:

$$RL = SL - 20 \log_{10}(r) \tag{5.2}$$

As a first approximation, spherical spreading is a good model for the propagation of terrestrial animal sounds produced in large open-air regions, such as grassland. Generally, if a bird sings on the ground up to about 10 m from a microphone, only spherical spreading needs to be considered. If the receiver is at a greater distance from the bird, then ground and atmospheric effects also must be considered. If the bird is flying overhead, then spherical spreading and atmospheric effects need to be considered when determining propagation characteristics.

If other sources of attenuation are negligible, then Eq. 5.2 can be used to calculate the source levels of a vocalizing animal located at distance $r$ from the receiver. For instance, if a bioacoustician measured $RL = 65$ dB re 20 μPa at a distance of 10 m from a singing bird, then $SL$ (at 1 m from the bird) becomes 65 dB re 20 μPa + 20 $\log_{10}(10)$ dB re 1 m = 85 dB re 20 μPa m (e.g., Dabelsteen 1981). Similarly, if somebody played back a sound at a known source level of 85 dB re 20 μPa m, then the predicted $RL$ at 1 km (= $10^3$ m) range would be 25 dB re 20 μPa, as $20 \log_{10}(10^3) = 60$.

In some environments, and for some sources (i.e., line sources rather than point sources), airborne sound propagation can be better described as cylindrical spreading. For an infinitely long line source, the propagation loss as a function of range becomes $PL_{cyl} = 10 \log_{10}(r)$ and so Eq. 5.1 becomes:

$$RL = SL - 10 \log_{10}(r) \tag{5.3}$$

Most biological line sources, however, are finite, such as a row of vocalizing birds on a power line. (Please be aware that this example is not a line source in the strict acoustic sense.) This means that geometrical spreading loss is somewhere between that of spherical and cylindrical spreading loss (Fig. 5.6). When the receiver distance from the finite line source is much less than the length of the finite line source, then the attenuation is close to that of an infinite line source (i.e., 10 $\log_{10}(r)$), whereas at distances comparable to or larger than the length of the finite line source, the latter acts more like a point source and attenuation develops as 20 $\log_{10}(r)$. At sufficiently long distances, all sources can be regarded as point sources.

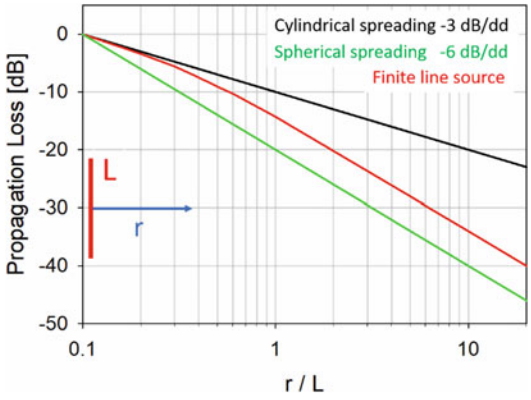



**Fig. 5.6** Propagation loss due to geometrical spreading in air from a finite length line source with distance $r$ relative to the length $L$ of the finite line source. At distances from the source shorter than $L$, the attenuation is close to 3 dB/dd (cylindrical attenuation), whereas at distances equal to or longer than $L$, the attenuation becomes 6 dB/dd (spherical attenuation); dd: distance doubled

The propagation loss, however, includes much more than geometrical spreading loss, since beyond some distance from the source, RL mostly becomes smaller with distance than predicted by Eqs. 5.2 or 5.3. To account for this extra attenuation, Marten and Marler (1977) introduced the term excess attenuation (EA). This includes a number of other effects such as atmospheric absorption, reflection and scattering, the ground effect, attenuation by vegetative cover, refraction by air temperature and wind gradients, and attenuation due to turbulence—and often there still is a rest attenuation not accounted for by these mechanisms (Wahlberg and Larsen 2017). While geometrical spreading is frequency-independent, most of the effects contributing to EA are frequency-dependent and thus alter the spectrum of the emitted sound.

In most bioacoustic scenarios, spherical attenuation applies, and Eq. 5.2 can be reformulated to:

$$RL = SL - 20\log_{10}(r) - EA \qquad (5.4)$$

The following sections investigate each of these components of EA.

### 5.2.3 Sound Absorption in Air

An important and predictable component of EA is attenuation by absorption in air. Absorption refers to the conversion of acoustic energy into heat, mostly due to molecular relaxation of air molecules and the air's shear viscosity. Absorption loss $EA_{abs}$ is directly proportional to the distance $r$ from the source:

$$EA_{abs} = \alpha r \qquad (5.5)$$

The absorption coefficient α (measured in dB/m) is a complex function of sound frequency, air temperature, relative humidity, and (to a lesser degree) atmospheric pressure (or altitude), in addition to characteristics of oxygen and nitrogen molecules (Attenborough 2007).

For instance, a 2-kHz signal propagating at standard atmospheric pressure (1 atm) and 20 °C is attenuated by about 0.9 dB/100 m, if the relative humidity (r.h.) is 60%, but by about 4.5 dB/100 m at 10% r.h. (Fig. 5.7). Generally, sound attenuation is greater in drier air than in damp, humid air. The effect is especially important at frequencies above 2 kHz. In other words, air acts as a low-pass filter enabling only low-frequency sound to travel over long distances from the source (Attenborough 2007; Wahlberg and Larsen 2017; Larsen and Radford 2018). Consequently, bats use high source levels to overcome the attenuation in air at high frequencies when they echolocate on targets at long distances. This low-pass filter effect is especially visible in the field for broadband sound signals produced by orthopterans and other insects (Römer 1998).

Sound absorption in air varies with time of day and season, mainly due to variations in the relative humidity, which usually peaks in the afternoon (see Larsson 2000; Attenborough 2007). So, if precise values of air absorption are needed in a field experiment, the relative humidity, atmospheric pressure, and air temperature must be measured over time and used in subsequent calculations (Wahlberg and Larsen 2017).

However, at the short distances (<100 m) where most acoustic communication between

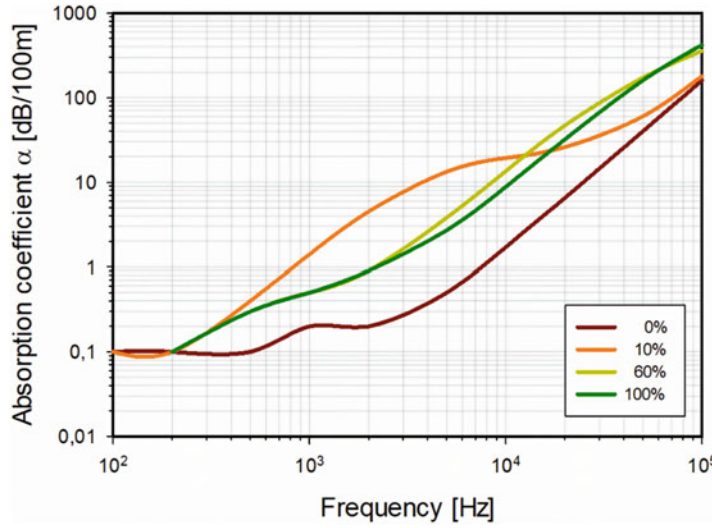



**Fig. 5.7** Sound absorption coefficients α in air (dB/100 m) at 20 °C versus frequency at four different relative humidities (r.h. %). Based on ISO 9613-1:1993 (International Organization for Standardization. ISO 9613-1:1993, Acoustics—Attenuation of sound during propagation outdoors—Part 1: Calculation of the absorption of sound by the atmosphere. International Organization for Standardization; https://www.iso.org/standard/17426.html; accessed 9 January 2021)

animals takes place and at frequencies below 10 kHz, the role of absorption in overall propagation loss is likely insignificant compared to other environmental factors. Garcia et al. (2012), for example, described the 40-Hz wing beat signals of drumming ruffed grouse (*Bonasa umbellus).* Theoretically, these sound signals would be reduced by 6 dB due to air absorption at a distance of 187 km from the drumming bird, whereas spherical spreading loss alone would have reduced the signal amplitudes to a level far below auditory threshold of most animals at a distance of 1 km already ($PL_{sph} = 60$ dB re 1 m).

### 5.2.4 Reflection, Scattering, and Diffraction

A second and less predictable component of EA is the attenuation caused by reflection, scattering, and diffraction. As a sound wave hits a hard surface, it is reflected. Reflection can be explained with Huygens' principle. In Fig. 5.8a, the rooster from Fig. 5.4a is very far away such that the wavefronts at any location appear planar (rather than circular) and the wave rays are parallel (rather than radial). Three incident rays are drawn, hitting the surface (e.g., a road) at times $t_1$, $t_2$, and $t_3$. By Huygens' principle, each point on the road that is hit acts as the source of a secondary wave. Two secondary wavefronts are shown at time $t_3$. From the time $t_1$, when the first ray hits, to the time $t_3$, the first wavefront has expanded quite a bit. The second wavefront was started at time $t_2$, when the second ray hit, and has expanded less by time $t_3$. The third ray is just starting its secondary wave at time $t_3$, with its secondary wavefront not yet visible. The tangent to the secondary wavefronts at time $t_3$ gives the new wavefront of the reflected wave. The angle of incidence (measured from the normal) is equal to the angle of reflection (also measured from the normal). This is referred to as the law of reflection. It applies to the so-called specular reflection (as from a mirror).

Reflection is not always specular but might instead be diffuse. In diffuse reflection, sound is scattered from the surface in all sorts of directions including the specular direction (Fig. 5.8b). This happens when the surface is not smooth but rough. Scattering depends on the ratio of the wavelength of sound to the size of the scatterer. When the sound wavelength is long (i.e., frequency is low) relative to the roughness of the surface, all the sound energy is reflected in the specular direction. When the wavelength is short (i.e., frequency is high) and less than the magnitude of the unevenness of the surface, then sound is scattered in other, non-specular directions. A gravel road, for instance, produces specular

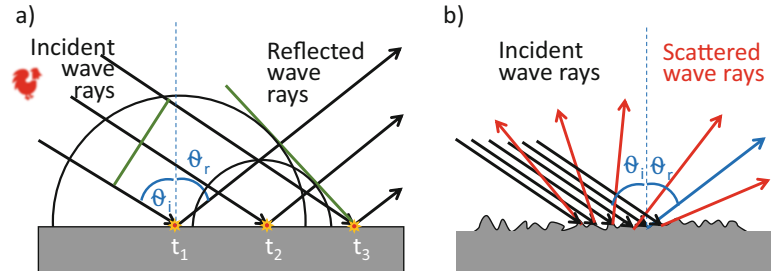



**Fig. 5.8** (**a**) Sketch of specular reflection of a plane wave (originating from a far-away rooster) off a hard surface. Wave fronts are shown as green lines; they are perpendicular to the wave rays, shown as black arrows. The three incident rays hit at times $t_1 - t_3$ at the locations marked by small suns. Each of these points creates a secondary wave by Huygens' principle. The secondary wavefronts superpose to yield the new wavefront of the reflected wave, shown at time $t_3$, when the third ray just hits, the second ray has started to grow a secondary wavefront, and the first ray has grown the largest wavefront. The angles of incidence $\theta_i$ are equal to the angles of reflection $\theta_r$. (**b**) Sketch of diffuse reflection off a rough surface where the unevenness is great compared to the wavelength of incident sound. While there is a reflected ray in the specular direction, too (indicated by a blue arrow), there are many other directions in which the incident sound is scattered (indicated by red arrows)

reflection at frequencies below 15–20 kHz, but at higher frequencies, where the gravel roughness is large relative to the wavelength, sound is scattered in different directions (Michelsen and Larsen 1983).

Reverberation is a result of multiple reflections and refers to the phenomenon of sound persisting even if the source is turned off. In canyons, caves, or other enclosures, sound bounces off the boundaries again and again. The reverberant sound field is the space that is dominated by reflected sound (as opposed to the field near the source where the direct sound dominates). Once the source is switched off, the reverberant field will continue to exist for some time, yet decay due to absorption by the medium, boundaries (e.g., the walls of a music room), and absorbers in the room (e.g., furniture and people). The more reflective the boundaries, the greater the reverberation.

Reverberation severely alters the structure of the received sound and is one of the least wanted effects in analysis of recorded animal sounds (Fig. 5.9). This type of signal degradation with propagation distance can be quantified by measuring the blur-ratio (see e.g., Dabelsteen et al. 1993). The received sound appears longer in duration than the emitted sound, with the delayed echoes forming a resulting "tail." This reverberation tail can be quantified as the tail-to-signal ratio (Holland et al. 2001). Consequently, leading edges of sound segments are relatively well-preserved, whereas ending edges are lost in reverberant environments.

Diffraction occurs when a sound wave is partially obstructed. In Fig. 5.10a, a plane wave (perhaps again from a far-away rooster) hits a wall with an opening in the center. The rays that hit the wall are reflected (not drawn). The rays that hit the opening pass straight through. By Huygens' principle, each point of the opening acts as a source of secondary waves. As the secondary wavefronts expand, they superpose to form new wavefronts that appear to bend behind the wall. This is termed diffraction. It also occurs when the obstruction is finite (Fig. 5.10b).

If the object that is in the path of a propagating sound wave becomes much smaller than a wall (e.g., a bush or maybe just an insect in the air), to the point where the wavelength is much greater (at least by a factor 10) than the size of the object, then the sound wave "ignores" the object and propagates without obstruction. The sound effectively cannot "see" the object; it is too small. In laboratory experiments, bioacousticians should therefore make sure that objects in the sound path from loudspeaker to experimental animal are at least 10 times smaller than the wavelength of the stimulus sound (Larsen 1995). When the

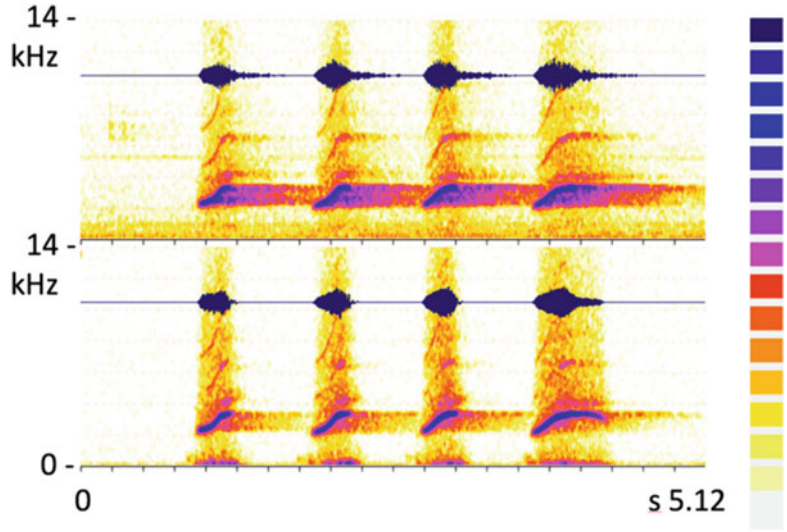



**Fig. 5.9** Spectrogram and envelope of a series of simple blackbird (*Turdus merula*) calls recorded at two different distances (amplitudes normalized and realigned in time). The spectrogram on top shows higher reverberation due to longer distance from the source than the bottom one. The color scale from white to black is 96 dB in 6-dB bins

wavelength is of the same order of magnitude as the object, or somewhat greater, then diffractive scattering occurs (Bradbury and Vehrencamp 2011). As the name suggests, this is a combination of diffraction and scattering, whereby some sound bends around the object and some sound scatters in all directions, leading to a complicated sound field.

Different surfaces or materials exhibit different degrees of sound reflection, absorption, and transmission. A hard, compact, smooth surface (such as a paved road, ice sheet, cave wall, canyon, subterranean tunnel, burrow wall, or wall of a captive animal's exhibit) reflects more and absorbs less acoustic energy than a porous, soft surface (such as tree leaves, grassy pastures, or forest canopy). Whether a surface or object is considered rough or smooth and hard or soft depends on the wavelength of the sound. In a mixed deciduous forest, reverberations for frequencies above 4 kHz are stronger with leaves on the trees than without leaves (Wiley and

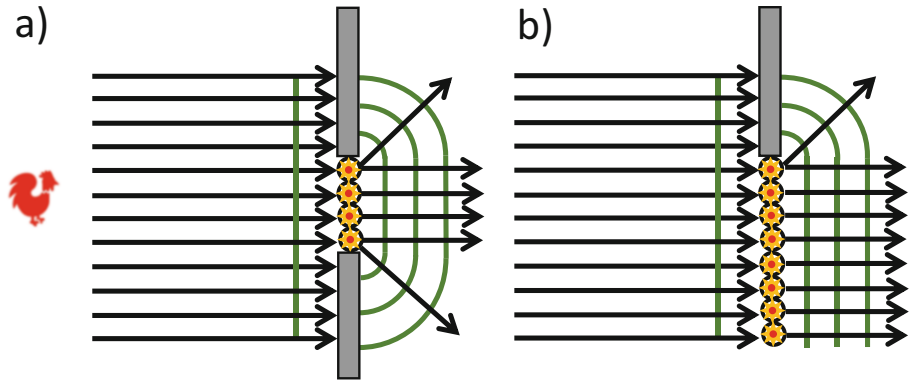



**Fig. 5.10** (**a**) Sketch of diffraction as a sound wave passes through an aperture. Wave rays are indicated by black arrows; wavefronts are indicated by green lines. As the plane wave from a distant rooster hits a wall, each point in the opening acts as a source (indicated by suns) of secondary waves. The secondary waves combine to create the new wavefronts shown at three successive instances in time. The wavefronts appear to bend behind the aperture. (**b**) Sketch of diffraction as a sound wave passes by a finite obstruction

Richards 1982). Reverberations essentially are absent in an open field on a calm day.

### 5.2.5 Ground Effect

Another component of EA is the so-called ground effect, which is always present in terrestrial sound propagation. The sound signal from a sender (S) located at some height above ground (e.g., a bird at 4 m) will reach a receiver (R; e.g., a recordist's microphone at 1.5 m) first by the direct path ($P_D$) and a moment later by the indirect and longer path when the signal has been reflected from the ground ($P_G$) (Fig. 5.11a). This results in a range-dependent interference pattern between the sound propagating along $P_D$ and $P_G$. The interference pattern has regions of enhanced received level (due to constructive interference) and of attenuated received level (due to destructive interference) at the position of R (Fig. 5.11b). The received sound signal is a distorted version of the emitted signal. It is said to be comb-filtered, as the destructive interference creates the "comb teeth" attenuating some frequencies in the signal, whereas the constructive interference enhances other frequencies of the signal. The magnitude of the ground effect depends on sound frequency, on geometry of the sender-receiver separation distance and height above ground, on the roughness and softness of the ground, and on atmospheric pressure, ambient temperature, relative

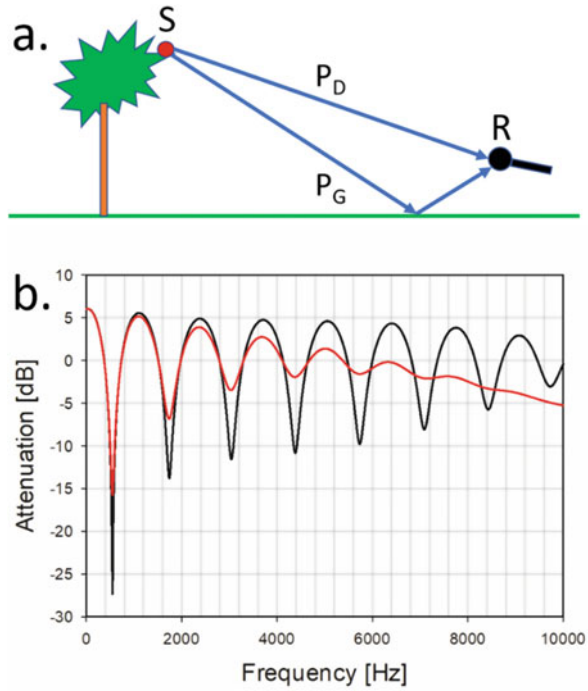



**Fig. 5.11** Predicted ground effect. (**a**) Sender 4 m above ground, Receiver 1.5 m above ground, horizontal separation distance 50 m (not to scale). The direct wave $P_D$ and the reflected wave $P_G$ superpose at R. (**b**) For frequencies whose wavelengths are in phase, superposition results in level enhancement up to 6 dB; at frequencies with wavelengths out of phase at R, levels are attenuated up to 20–30 dB. Black curve: The curve represents the predicted decibel values that need to be added to the geometric attenuation loss. The ground was modeled as a grass-covered field (flow resistivity 100 kPa s $m^{-2}$, porosity 30%, layer depth 0.01 m). Red curve: As in the black curve, but more realistic air absorption (at 20 °C, 75% relative humidity, standard atmospheric pressure) and moderate turbulence (mean-squared refractive index of $10^{-5}$) were added. Effects of temperature and wind-induced refraction were excluded in the model, which was developed by Keith Attenborough and Shahram Taherzadeh and improved by Kenneth Kragh Jensen

humidity, and turbulence (see Attenborough et al. 2007). Acoustically hard ground surfaces (such as rock or consolidated sand) produce comb-filter effects over a wide frequency range extending to relatively high frequencies, whereas acoustically soft surfaces (such as grasslands, forest floors, or unpacked snow) mainly generate the ground effect at low frequencies. Recordists may reduce the ground effect by placing microphones as high as practically possible above soft ground. For a general introduction to the phenomenon, see Michelsen and Larsen (1983) or Wahlberg and Larsen (2017). For a comparison between ground effect models and outdoor recordings, see Jensen et al. (2008).

### 5.2.6 Attenuation by Vegetative Cover

Absorption of sound by vegetation is a component of EA that can further dissipate airborne sounds over distance as acoustic energy is converted to heat in the plant material by viscous friction. The absorption of sound in vegetation depends on the material composition and hardness of the surfaces including the soft ground often found especially in woodland. Leaves absorb more sound energy than a tree trunk; whereas a tree trunk reflects more sound than leaves do. All of this is frequency-dependent.

This component of EA obeys no simple rules and needs to be measured by propagation experiments in the field (e.g., Dabelsteen et al. 1993). Aylor (1972a, b) measured sound propagation loss through various crops, bushes, and trees by broadcasting from a loudspeaker and recording at some distance with a microphone. He found foliage enhanced absorption and scattering. Price et al. (1988) modeled and measured attenuation by vegetation in different forest environments and documented scattering from tree trunks, enhanced ground effect in the presence of mature forest litter, and attenuation by foliage. Foliage attenuation had the greatest effect above 1 kHz and increased almost linearly with the logarithm of frequency. Through mixed coniferous forest, for instance, the attenuation over 24 m varied from about 5 dB at 2 kHz to 10 dB at 4 kHz, which is the range of dominant frequencies in many songbird songs. This foliage attenuation is less than, but needs to be added to, the 28-dB attenuation caused by spherical spreading over the same distance (Eq. 5.2).

Some research on sound propagation through vegetation was motivated by a desire to attenuate anthropogenic noise such as road noise, but generally and most surprisingly dense foliage only accounts for a small amount of attenuation. Martínez-Sala et al. (2006) concluded that a 15-m wide patch of regularly spaced trees could attenuate car noise by at least 6 dB. The effect was similar for more traditional noise barriers. Defrance et al. (2002), for instance, found that a 100-m wide forest strip was effective at providing an acoustical barrier to noise, such as shown in Fig. 5.12, where octave-band sound was broadcast through dense foliage and recorded at different distances in the forest.

At present, vegetation attenuation is not well understood. A much larger database is needed before it is possible to accurately predict the effect of different kinds of vegetation on sound propagation (see Attenborough et al. 2007).

### 5.2.7 Speed of Sound in Still Air

The speed of sound in still air is affected only by the ambient air temperature and, to a minimal extent, air pressure (or altitude). If the sound propagates under windy conditions, however, the effective speed of sound will be modified by the wind velocity such that the wind velocity of a tailwind will add to the speed of sound and the wind velocity of a headwind will subtract from the speed of sound.

The speed of sound determines the arrival time of a signal from the sender to the receiver and bends a propagating sound wave away from higher air temperature and towards lower air temperature (or from higher wind velocity towards lower wind velocity). The speed of sound in air at 21 °C is 344 m/s. At freezing point, 0 °C, the speed of sound in air is 331 m/s. A good

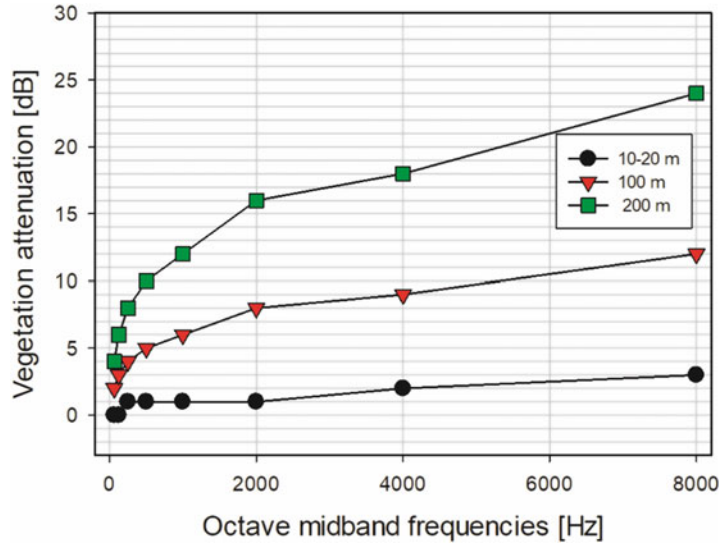



**Fig. 5.12** Attenuation of octave bands of noise (63 Hz to 8000 Hz) after propagating three distances through dense foliage. Data from ISO 9613-2:1996 (International Organization for Standardization. ISO 9613-2:1996, Acoustics—Attenuation of sound during propagation outdoors—Part 2: General method of calculation. International Organization for Standardization; https://www.iso.org/standard/20649.html; accessed 9 January 2021)

approximation of the speed of sound $c$ in dry air with 0.04% $CO_2$ and temperature $T_c$ (in °C) is:

$$c = (331.45 + 0.607\, T_c)\ \mathrm{m/s} \tag{5.6}$$

### 5.2.8 Refraction by Air Temperature Gradients in Still Air

Refraction is the change of the direction of sound propagation due to changes in the speed of sound. In the example of Fig. 5.13a, a plane wave in medium 1 hits an interface with medium 2. Some of the acoustic energy might be reflected (as in Fig. 5.8a, not drawn in Fig. 5.13a), and some of the energy is transmitted. The transmitted wave is refracted, because the speeds of sound differ in the two media. If $c_1 > c_2$, then the transmitted wave bends towards the normal (i.e., away from the interface; Fig. 5.13a); if $c_1 < c_2$, then the transmitted wave bends away from the normal (i.e., towards the interface; Fig. 5.13b). The angles of incidence and refraction (transmission) are related via Snell's law (named after Dutch astronomer and mathematician Willebrord Snell):

$$\frac{\sin\theta_i}{\sin\theta_t} = \frac{c_1}{c_2} \tag{5.7}$$

Note that, while the frequency of the sound does not change during transmission, the wavelength does change. With $c = \lambda f$ (see Chap. 4, section on the speed of sound), the wavelength is smaller in the medium with lower sound speed.

Refraction of sound waves in air is a common phenomenon due to vertical gradients of air temperature and/or wind velocity. A gradual change in sound speed is illustrated in Fig. 5.13b, where the rays bend more and more upwards as the sound speed increases. In terrestrial environments, the sound source is typically located close to the ground. A sound speed profile that has the speed of sound increase with altitude is downward refracting, while a sound speed profile that has the speed of sound decrease with altitude is upward refracting. Bent propagation paths have the effect that sound appears to arrive from a non-intuitive (i.e., not straight-line) direction. This phenomenon is like an acoustic mirage in analogy to optical mirages, which produce displaced images of far-away objects and which are also caused by refraction (of light).

The EA from refraction may be positive or negative, and so RL may be smaller or greater

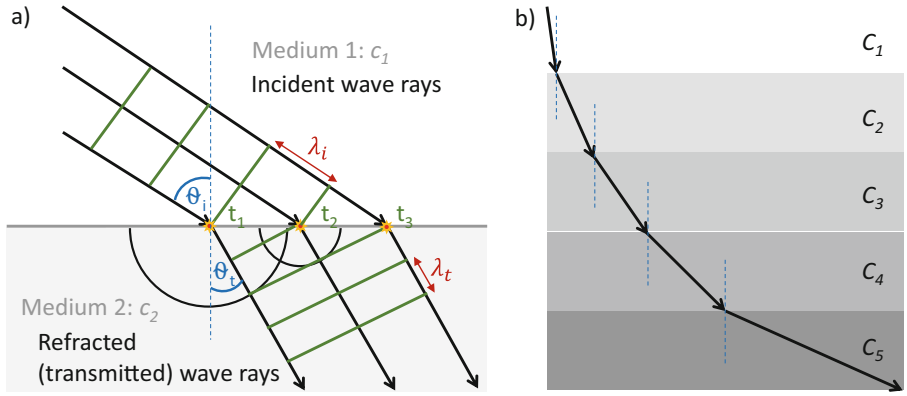



**Fig. 5.13** (**a**) Sketch of refraction at a boundary between medium 1 (high sound speed) and medium 2 (low sound speed). Three rays (black arrows) are shown, hitting the interface at times $t_1$-$t_3$. Each gives rise to secondary waves (by Huygens' principle) starting at the points marked with small suns. At time $t_3$, the third ray just meets the interface, the second ray has produced a small secondary wave, and the first ray's secondary wave has grown quite a bit. Drawing the tangent to the secondary waves at time $t_3$ yields the new wavefront (green line) in the second medium. With rays, by definition, being perpendicular to the wavefronts, it can be seen that the rays bend towards the normal in the second medium ($\theta_t < \theta_i$). Successive wavefronts are drawn to show that they are spaced farther apart in the medium with higher sound speed, and so the wavelength $\lambda$ is greater in the medium with higher sound speed. (**b**) Sketch of gradual refraction by a vertical gradient in sound speed. In the illustrated example, $c_1 < c_2 < c_3 < c_4 < c_5$

than predicted without a refracting atmosphere. Air temperature varies throughout the day and creates varying temperature gradients. So, recording at the same location at a different time of day can produce different results. Therefore, taking periodic measurements of the ambient temperature at different heights above the ground can provide the researcher with a notion of whether sound propagation is changing and at what pace.

In still air during daytime, the air is both warmer and more humid close to the ground and a stable air temperature gradient can be established with warmer air near the ground, because of sunlight heating the ground, which warms up much faster than the overlaying air. At higher elevations, the air temperature decreases by 0.01 °C/m (Fig. 5.14a). Sound waves consequently bend away from locations near the ground where the temperature is higher and upwards towards locations with lower temperatures (Fig. 5.14b). Horizontal rays will be directed upwards as will downwards directed rays after bouncing from the ground. Therefore, a certain limiting ray exists that defines a shadow zone around the sound source, where the sound level decreases way faster than predicted from distance alone (Fig. 5.14b). While the shadow zone cannot be reached by a direct path, it may be ensonified by reflection off houses (or other reflectors) in the vicinity and by paths passing through turbulence, and the shadow zone is thus not totally quiet.

For example, on a sunny day with little wind, the air temperature can be 30 °C at the ground ($c = 351$ m/s), but at 2–3 m above ground, the temperature may be only 25 °C ($c = 347$ m/s). This decrease continues up through the atmosphere by 1 °C/100 m, the so-called temperature lapse. With such an air temperature gradient, the sound rays from a sound source located a few meters above ground will bend upwards, because part of the wave closest to the warmer ground will travel the fastest. In a carefully conducted experiment, a combination of upward refraction, strong upwind propagation, and air absorption was measured to reduce the level of propagating sound at a distance of 640 m by up to 20 dB more than predicted from Eq. 5.2 (Attenborough

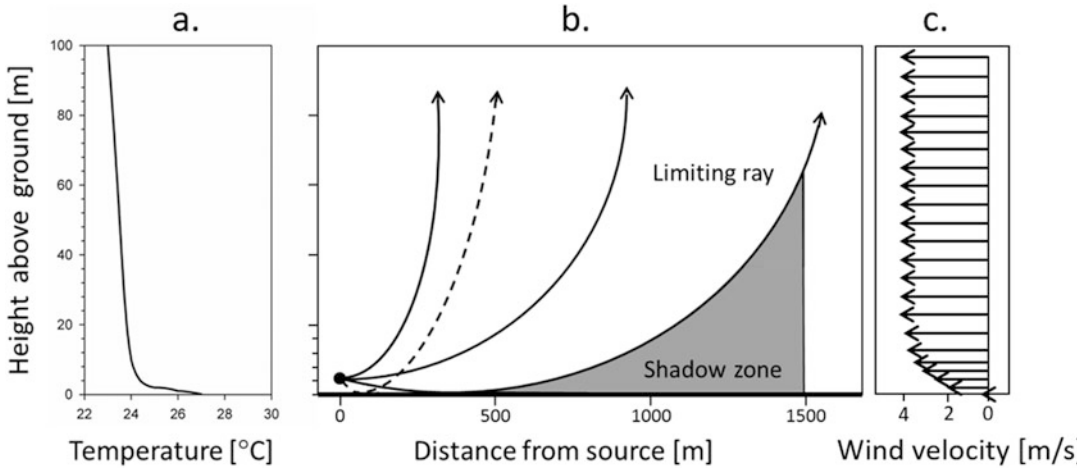



**Fig. 5.14** Sketch of the effects of upward refracting sound speed gradients on outdoor sound propagation. (**a**) Temperature profile: Air temperature and consequently sound speed increases towards the ground in still air. (**b**) Ray traces: Sounds from a source (filled circle, here 5 m above ground) are refracted upwards, creating a circular shadow zone close to the ground around the source. Dashed line indicates a sound ray bouncing off the ground. (**c**) Wind velocity profile: Similar upward refraction is created upwind. Arrows indicate wind direction towards the source ("headwind") and their length wind speed. Reprinted by permission from Springer Nature. Acoustic Conditions Affecting Sound Communication in Air and Underwater, Larsen and Radford (2018), Fig. 5.5.4. In: H Slabbekoorn, RJ Dooling, AN Popper and RR Fay (eds). Effects of Anthropogenic Noise on Animals, Springer Handbook of Acoustic Research 66, Springer Science and Business Media, LLC, part of Springer Nature: New York, Heidelberg, Dordrecht, London. pp. 109–144. https://doi.org/10.1007/978-1-4939-8574-6_5. © Springer Nature, 2018. All rights reserved

2007). Perhaps for this reason, birds do not commonly sing in open environments near the ground on sunny days. Rather, they sing in flight well above ground, or from a perch (Wiley 2009).

On calm nights, the opposite air temperature gradient can occur close to ground (called temperature inversion) as it cools faster than the overlaying air. Air temperatures increase up to 50–100 m above ground before decreasing again with altitude. Therefore, sound rays bend downwards and hit the ground (Fig. 5.15). A temperature inversion favors long-distance sound propagation as it leads to higher received levels than predicted by spherical spreading. For this reason, nocturnal communication distances of low-frequency African savanna elephant (*Loxodonta africana*) sound doubled on the savanna to as much as 10 km (Garstang et al. 1995). In these conditions, sound energy is channeled making spreading losses effectively cylindrical, rather than spherical within the surface layer. Garstang (2010) suggested that a loud infrasonic elephant call during the middle of the day would travel no more than 1 km (i.e., be heard over an area of 3 $km^2$), but an elephant call at night might be heard over an area of 300 $km^2$ (see also, Garstang et al. 1995; Larom et al. 1997). Elephants might adjust timing and abundance of their low-frequency calls and apply them specifically for long-distance communication according to atmospheric conditions.

An air temperature gradient can arise in other locations than just close to ground. Geiger (1965) found the air in and above the forest canopy beginning to warm immediately after sunrise, whereas the air below the canopy was slower to respond. This creates a bilinear sound speed profile with an upward refracting gradient above the canopy and a downward refracting gradient below the canopy. So, for a short period after sunrise, vocalizing birds and, for instance, howler monkeys (*Alouatta* sp.) located below the canopy can increase the range of their vocalizations relative to later in the day (Wiley and Richards 1978; Wiley 2009).

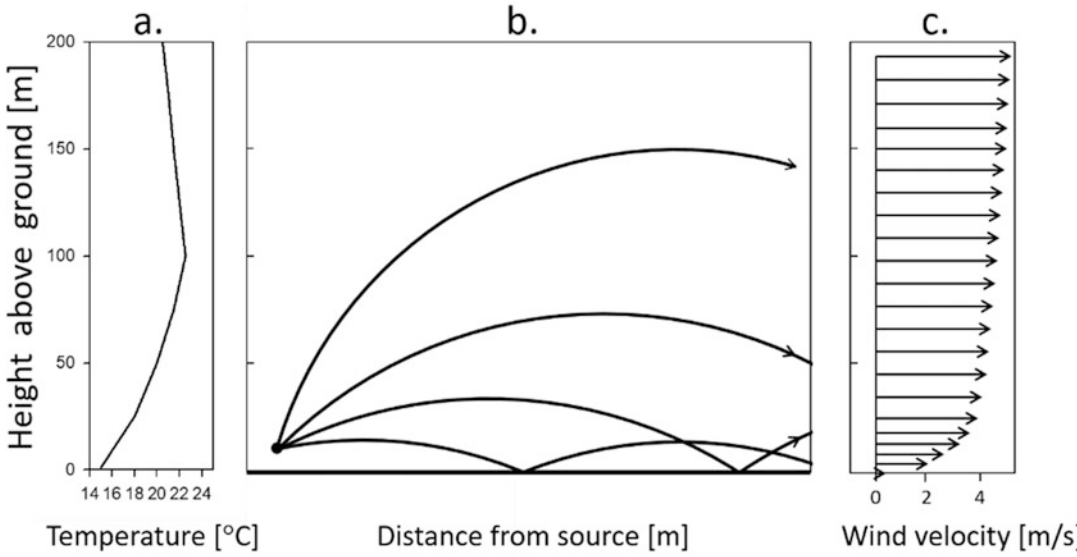



**Fig. 5.15** Sketch of the effects of downward refracting sound speed gradients on outdoor sound propagation. (**a**) Temperature profile: On calm nights, air temperature and consequently sound speed may increase with height above ground until temperature lapse starts. (**b**) Ray traces: Sounds from a source (filled circle, here 5–10 m above ground) are refracted downwards, creating higher sound levels with distance than predicted from spherical spreading. (**c**) Wind velocity profile: Similar downward refraction with increased sound levels may be created downwind. Arrows indicate wind direction away from the source ("tailwind") and their length wind speed. Reprinted by permission from Springer Nature. Acoustic Conditions Affecting Sound Communication in Air and Underwater, Larsen and Radford (2018), Fig. 5.5.5. In: H Slabbekoorn, RJ Dooling, AN Popper and RR Fay (eds). Effects of Anthropogenic Noise on Animals, Springer Handbook of Acoustic Research 66, Springer Science and Business Media, LLC, part of Springer Nature: New York, Heidelberg, Dordrecht, London. pp. 109–144. https://doi.org/10.1007/978-1-4939-8574-6_5. © Springer Nature, 2018. All rights reserved

### 5.2.9 Refraction by Gradients of Wind Velocity

Strong air temperature gradients cannot exist during strong wind conditions, so the effects of wind velocity on sound propagation in open environments are more influential than air temperature gradients (Attenborough 2007). Wind may cause a shift in sound direction such that the appearance from where the sound is generated differs from where it is actually sent (acoustic mirage). Wind velocity gradients can enhance or impede sound propagation, leading to negative or positive EA. The actual speed of sound is the sum of the air temperature-generated speed of sound and the net wind velocity.

Attenborough et al. (2007) reported the general relationship between the sound speed profile $c(z)$, the air temperature profile $T(z)$, and the wind velocity profile $u(z)$, where $z$ is the height above ground, when the wind blows in the direction of sound propagation (when the wind blows against propagation, $-u(z)$ is added):

$$c(z) = c(0)\sqrt{\frac{T(z) + 273.15}{273.15}} + u(z) \qquad (5.8)$$

Wind velocity is lowest at the ground and increases with altitude (Figs. 5.14c, 5.15c). Sound traveling upwind refracts upwards and sound traveling downwind refracts downward (Fig. 5.14b, Fig. 5.15b). As with temperature gradients, this creates a shadow zone upwind (Fig. 5.14b), where the sound is not heard. Downwind, sounds propagate in a channeled way (Fig. 5.15b) with less loss. Sound attenuates more against the wind than with the wind. Despite this

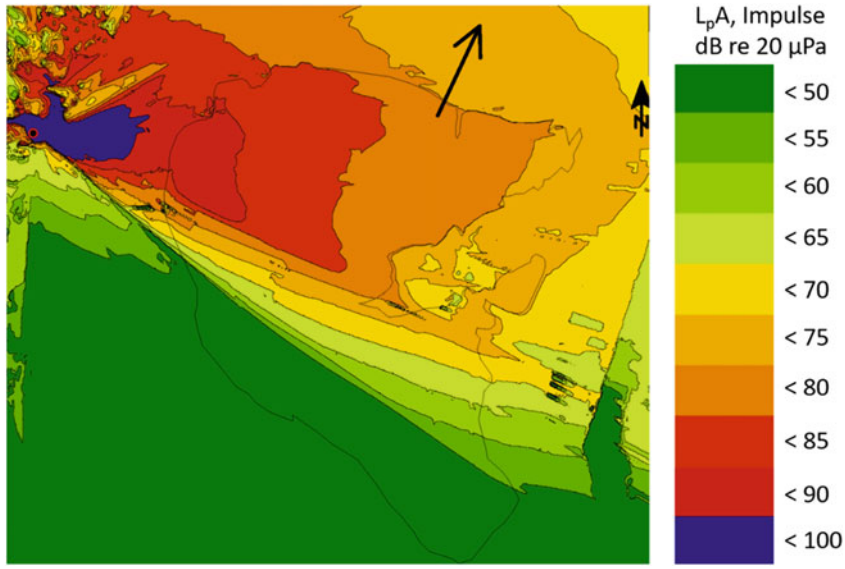



**Fig. 5.16** Noise map showing the received levels 50 cm above ground of a gunshot fired towards east at a location (small red circle in dark blue area upper left corner) close to a lake (lake contour lines indicated by thin black curves) with varied topography. The color coding indicates iso-dB-curves in 5-dB steps. The dark arrow indicates wind direction and its length corresponds to 300 m on the ground. Note how the wind attenuates the gunshot upwind and enhances it downwind. Noise map calculated by DELTA—a part of FORCE Technology, Hørsholm, Denmark, using *Nord2000* software (https://eng.mst.dk/air-noise-waste/noise/traffic-noise/nord2000-nordic-noise-prediction-method/; accessed 23 December 2020). Figure donated by Jesper Madsen, Aarhus University

common phenomenon, Wiley (2009) commented that there are no documented cases of animals selectively communicating downwind. But refraction by gradients of wind velocity played a significant role in Civil War battles in the rolling hills of the eastern U.S. There was no radio communication in the nineteenth century, so commanders often depended on what they heard of the battle in front of them to make decisions about troop movements. An acoustic shadow zone existed during the Battle of Gettysburg and commanders could not hear the sounds of battle just 10 miles away, whereas people 150 miles away in Pittsburgh clearly heard the skirmish (Ross 2000).

Sound maps portray the attenuation of sound over distance from a source. The maps take a bird's-eye view, showing attenuation in 360° about a sound source. Such maps can be produced at a specific receiver altitude, or commonly show maximum received levels over a range of altitudes with the intent of yielding "conservative" estimates of received level. The attenuation pattern radiating from the sound source is typically irregular in shape (rather than concentric) and helps identify environmental conditions that impede or promote sound propagation. Sound mapping tools can commonly utilize data on topography and ground absorption, air temperature, and wind direction and speed. The example in Fig. 5.16 shows how wind attenuated noise from a gunshot upwind but enhanced received levels downwind.

### 5.2.10 Attenuation from Air Turbulence

Turbulence refers to unsteady and irregular motion of the air. It is very difficult to model and predict. It may be mechanically or thermally induced. Mechanical turbulence is caused by friction, for example, when air moves over rough ground or past obstacles such as houses and trees. Friction causes eddies and thus turbulence.

This turbulence is stronger in higher wind speeds and rougher terrain. Turbulence is particularly great during fall winds, which shoot down the slope of a mountain. Thermal turbulence is created when the sun heats the ground unevenly. For example, bare ground warms up faster than fields with vegetative cover or bodies of water. Convective air currents are established with warm and less dense air rising and cold and denser air sinking. These currents, in turn, may generate eddies. Eddies may extend from the ground to a few hundred meters height. They can be of various sizes (height and diameter) and larger eddies may break up into smaller ones. Because of air temperature, gradients and wind, air is always in motion and this motion may always generate turbulence.

Turbulence causes EA, which increases with distance from the source, with the level of turbulence, and with sound frequency (see red curve in Fig. 5.11b). EA is typically highest during daytime and on hot sunny days. A characteristic of turbulence on sound propagation is that received levels at a fixed location quickly fluctuate with time and, at some range, this fluctuation stabilizes at a standard deviation of about 6 dB (Daigle et al. 1983). Van Staaden and Römer (1997), for instance, reported that at night, the sound pressure level of the song of an African bladder grasshopper (*Bullacris intermedia*) over open grassland was reduced with distance very close to the expected 6-dB per doubling of distance of spherical attenuation. However, during daytime, the attenuation was much larger and more variable due to air turbulence.

For more in-depth reading on outdoor sound propagation, please see Attenborough et al. (2007), Attenborough et al. (2007), Larsen and Wahlberg (2017), Wahlberg and Larsen (2017), or Larsen and Radford (2018).

## 5.3 The Source-Path-Receiver Model for Animal Acoustic Communication

The SPRM can be used to examine acoustic communication among animals. In the example of Fig. 5.17, two gentoo penguins (*Pygoscelis papua*) are communicating within their nesting colony in Antarctica. The sender (i.e., the source) emits a penguin display call. The call spreads through the habitat, experiencing various forms of attenuation. The receiver is another gentoo penguin. It might respond acoustically and thus become the next sender. Whether this two-way acoustic communication is successful, depends on a number of parameters.

The locations of sender and receiver matter; the closer together they are, the better the communication—most likely. If the source emission pattern is directional rather than omnidirectional (i.e., the call can be emitted in a specific direction), then the orientation of the sender towards the receiver matters. Similarly, if the receiver's hearing is directional, then the receiver's orientation affects communication success. A stronger source level will increase the likelihood of successful reception, unless the environment is highly reverberant, in which case the echoes would also be louder and potentially interfere with communication success. The frequency content of the call matters, because different frequencies propagate differently, and the hearing abilities of the receiver are frequency-dependent.

Along the path, some of the call energy is lost due to geometrical spreading and some is absorbed by the air, snow, and soil. The direction of propagation changes due to reflection and scattering off rocks, and due to refraction by sound speed gradients in air. Diffraction around mountains might play a role over longer ranges. Ambient noise in the environment does not affect sound propagation; i.e., it neither leads to attenuation nor changes the direction of propagation.

Ambient noise in the environment affects whether the call is received and correctly interpreted. Ambient noise can be of abiotic, biotic, or anthropogenic origin. Wind causes noise, as do waves and breaking ice. The other penguins in the colony create ambient noise with their own acoustic communications. Human presence (e.g., chatting tourists stomping through the snow towards the penguin colony) might add to the ambient noise. Ambient noise at the location of the receiver lowers the signal-to-noise ratio

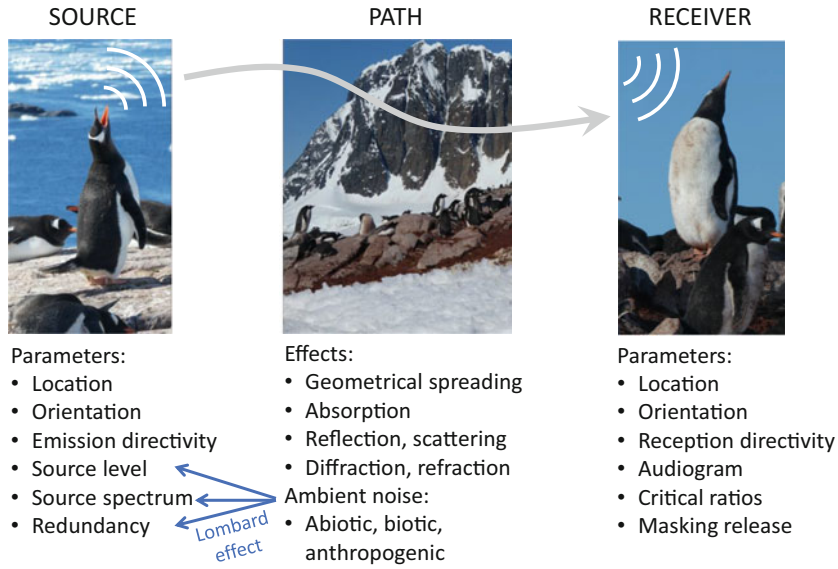



**Fig. 5.17** Example of the SPRM for animal acoustic communication. The source is a gentoo penguin emitting its display call within its nesting colony in Antarctica. The sound propagation path takes the call through the local habitat. The receiver is another gentoo penguin in a neighboring colony who might respond acoustically, thereby becoming the next source. The parameters that affect successful communication are listed below the source and the receiver. Along the path, the call experiences various propagation effects leading to attenuation. Ambient noise in the habitat stems from waves, wind, and ice (abiotic), other penguins (biotic), and perhaps humans (anthropogenic). Ambient noise at the receiver reduces the signal-to-noise ratio and hence the detectability of the call. Ambient noise at the source may lead to increases in source level and repetition (redundancy) and shifts in spectral content (Lombard effect)

(SNR) at which the call is received. The critical ratios (specific to the receiver's auditory system; see Chap. 10) dictate, below which SNR the call is masked by the ambient noise and thus not detected. At intermediate SNRs, the call might be detected, but not correctly interpreted. Masking-release processes (also specific to the receiver's auditory system) include comodulation masking release and spatial release from masking (e.g., Erbe et al. 2016) and aid signal detection and interpretation. Ambient noise at the sender may lead to the Lombard effect (Lombard 1911), whereby the sender raises the source level of its call, actively changes the spectral characteristics to move sound energy out of the frequency band most at risk from masking, and repeats the call to increase the likelihood of reception. Finally, ambient noise may instill anti-masking strategies in both sender and receiver whereby they change their location and orientation (both towards each other) to foster communication success.

### 5.3.1 The Sender

In animal acoustic communication, the signal that is being sent depends on the sender's species, demographic parameters, behavioral state, and many other factors. Obviously, different taxonomic groups produce different sounds, ranging from infrasonic rumbles of elephants to ultrasonic clicks of bats (see Chap. 8 on classifying animal sounds). But even closely-related species may be told apart acoustically. For example, Gerhardt (1991) found that the number of pulses in the advertisement call in male Eastern gray treefrogs (*Dryophytes versicolor*) and Cope's gray treefrogs (*Dryophytes chrysoscelis*) is the major cue distinguishing sympatric males who are similar in size and color. While species-specific calls of bats have been recognized for decades (Balcombe and Fenton 1988; Fenton and Bell 1981; O'Farrell et al. 1999), more recently, acoustic differences have been noted in bat

species that are difficult to tell apart morphologically (Gannon et al. 2001; Gannon et al. 2003; Gannon and Racz 2006). The more we record and document species' repertoires, the more successful bioacousticians will become at identifying the sender's species.

Within the same species, populations living in different geographic regions and habitats may exhibit differences in their sounds, as demonstrated for Italian vs. English tawny owls (*Strix aluco*; Galeotti et al. 1996), pikas (*Ochotona* spp.; Trefry and Hik 2010), and chimpanzees (*Pan troglodytes schweinfurthii*; Mitani et al. 1992). Animals can tell conspecifics from a different region or population apart. Auditory neighbor-stranger discrimination has been demonstrated, for instance, in concave-eared torrent frogs (*Odorrana tormota*; Feng et al. 2009) and alder flycatchers (*Empidonax alnorum*; Lovell and Lein 2004), where territory holders respond less aggressively towards played-back neighbor songs than to those of strangers, the "dear enemy effect."

Not just population identity, but even individual identity may be encoded in the outgoing signal; for example, in oilbirds (*Steatornis caripensis*; Suthers 1994), banded mongoose (*Mungos mungo*; Fig. 5.18; Jansen et al. 2012), and in fallow deer (*Dama dama*; Vannoni and McEligott 2007). Galeotti and Pavan (1991) studied an urban population of non-songbirds, tawny owls, in Pavia, Italy, and demonstrated that the males' territorial hoots have a clear species-specific structure with individual variations mainly in the final note of the call. Bats use individualized calls as they aggregate. For example, Melendez and Feng (2010) determined that communication calls of little brown bats (*Myotis lucifugus*) were individually distinct in minimum and maximum frequency, and call duration. Individual pallid bats (*Antrozous pallidus*) emitted unique calls below the frequency of their echolocation clicks and in the presence of other bats (Arnold and Wilkinson 2011). Wilkinson and Boughman (1998) provided evidence that the greater spear-nosed bat (*Phyllostomus hastatus*) used individual social calls to coordinate feeding on clumped nectar and fruit resources. Colonial

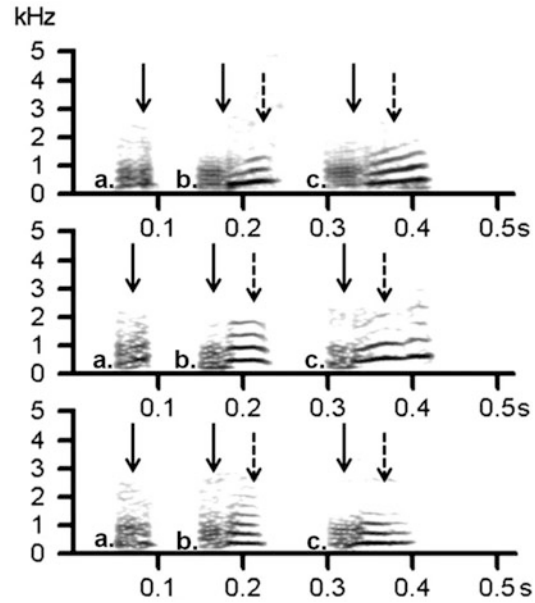



**Fig. 5.18** Spectrograms of close calls of three banded mongoose (two females and one male; top to bottom) during a. digging, b. searching, and c. moving between foraging sites. Black arrows point to the individually stable foundation of each call. Dashed arrows point to the harmonic extension, the duration of which was correlated with behavior (Jansen et al. 2012). © Jansen et al.; https://link.springer.com/article/10.1186/1741-7007-10-97. Published under a Creative Commons Attribution License; https://creativecommons.org/licenses/by/2.0/

animals, such as penguins, gulls, pinnipeds, and bats especially rely on individual acoustic recognition between a mother and offspring. These mothers often leave their young in a colony while they forage, so proper recognition of their own young upon return is important to fitness. Especially in birds without nests and physical landmarks such as king penguins (*Aptenodytes patagonicus*), acoustic recognition between parents and chicks becomes critical (Aubin and Jouventin 2002; Searby et al. 2004).

As organisms grow, their physical dimensions and size of their sound-producing organs become larger. Generally, emitted sounds transition from high-frequency, low-amplitude sounds to low-frequency, high-amplitude sounds (Hardouin et al. 2014). It is partly a consequence of the simple physiology that animals cannot efficiently emit sounds with wavelengths longer than the dimensions of their sound-emitting organs (e.g.,

see Michelsen 1992; Genevois and Bretagnolle 1994; Fletcher 2004, and Larsen and Wahlberg 2017). For instance, Charlton et al. (2011) reported that increased body size in male koalas (*Phascolarctos cinereus*) was reflected in the closer spacing of vocalization formants. (Formants refer to a concentration of acoustic energy around particular frequencies caused by resonances in the vocal tract.) Stoeger-Horwath et al. (2007) reported age-dependent variations in the grunt and trumpet calls of African savanna elephants. The grunts were only recorded in individuals less than 2 months of age and infants never produced trumpet calls until they were 3 months old. The authors also reported age-dependent variations in the low-frequency rumble; older individuals rumbled at a lower fundamental frequency than younger individuals, and there also was a tendency for rumble duration to increase slightly with age. Weddell seal (*Leptonychotes weddellii*) pups on rookeries emit high-frequency calls that transition into low-frequency adult calls used exclusively while hauled-out on the ice (Thomas and Kuechle 1982). Reby and McComb (2003) reported that lower-frequency male roars in red deer (*Cervus elaphus*) stags were associated with greater age and weight, so provided "honest" cues about reproductive condition.

In many species, sex-specific differences in the acoustic repertoires are employed to insure proper mate selection (Hardouin et al. 2014). The sender's reproductive state and drive for mating often is represented in its acoustic signals. In songbirds and many orthopteran insects, only males sing (Miller et al. 2007; Riede et al. 2010). Songs are under the influence of reproductive hormones associated with courtship, and songbird songs are long, complex, and repeated in a typical and recognizable sequence of sounds. In species in which males compete acoustically to attract a female mate, a substandard mating call could indicate immaturity, agedness, or poor health of the caller. For example, Hardouin et al. (2007) examined hoots by 17 male scops owls (*Otus scops*) on the Isle of Oléron, France. Heavier male owls made lower-frequency hoots, which could give them a competitive mating advantage over lighter weight males.

Context further determines acoustic signaling. For example, predators often hunt quietly, and prey remain silent when it is aware of being stalked. A classic case where (prey) moths attempt to jam (predator) bat echolocation signals with a counter signal to confuse the approaching predator has developed another twist. Ter Hofstede and Ratcliffe (2016) found that, "specific predator counter-adaptations include calling at frequencies outside the sensitivity range of most eared prey, changing the pattern and frequency of echolocation calls during prey pursuit, and quiet, or 'stealth,' echolocation." Acoustic interactions between a parent and offspring are often brief and relatively quiet to conceal and protect the young. In contrast, messages with a high reproductive value, such as mating calls or territorial defense calls, and calls with high survival value, such as infant distress calls or adult alarm calls, are produced loudly and repeatedly. To this point, it has been shown that distress calls of three species of pipistrelle bats (*Pipistrellus nathusii*, *P. pipistrellus*, and *P. pygmaeus*) were structurally convergent, "consisting of a series of downward-sweeping, frequency-modulated elements of short duration and high intensity with a relatively strong harmonic content" (Russ et al. 2004). The study suggested that it was not as important to have species-specific signals as it was to have some device that produced a mobbing by bats of the predator regardless of species of bat.

Ambient noise at the location of the sender may also affect signal emission level, repetition, and spectral shifts (collectively called the Lombard effect; Brumm and Zollinger 2011). For instance, male túngara frogs (*Engystomops pustulosus*) increased the level, repetition, and complexity of their calls when noise overlapped with their normal frequency band of calling but not when noise was higher and non-overlapping in frequency (Halfwerk et al. 2016). Brumm (2004) and Brumm and Todt (2003) noted that birds in a noisy environment called louder and more often, and repositioned themselves, possibly to increase the likelihood of the sound being received. Similarly, greater horseshoe bats (*Rhinolophus ferrumequinum*) increased their call level and shifted frequency in noisy

environments (Hage et al. 2013). Eliades and Wang (2012) examined the neural processes underlying the Lombard effect in marmoset monkeys (*Callithrix jacchus*) and found that increased vocal intensity was accompanied by a change in auditory cortex activity toward neural response patterns observed during vocalizations under normal feedback conditions.

Many animal communication calls are close to being omnidirectional, radiating equally in all directions—at least at their lower frequencies (Larsen and Dabelsteen 1990). However, some bird species (e.g., juncos, warblers, and finches) showed an ability to focus their calls in the direction of an owl to warn-off the predator. Yorzinski and Patricelli (2009) examined the acoustic directionality of antipredator calls of 10 species of passerines and found that some birds would "call out of the side of their beaks" with their head pointed away from conspecifics in an apparent attempt at ventriloquist behavior. Whether terrestrial animals can actively change the sound emission directivity in response to noise (in order to enhance acoustic communication) needs to be investigated.

### 5.3.2 The Path and the Acoustic Environment

As the signal leaves the sender and travels through the environment, it is subjected to various forms of attenuation (as detailed above) and so the level at the receiver location is less than the source level. In addition, ambient noise at the receiver location reduces the SNR, making it harder for the receiver to detect the signal. Ambient noise may be classed according to its sources: abiotic, biotic, or anthropogenic. Chapter 7 provides a detailed overview of ambient noise with example spectrograms.

In terms of abiotic ambient noise, wind is a major contributor and its noise level increases with wind speed. In addition, remember that the direction of wind (i.e., upwind or downwind) affects the distance that sounds propagate. Wind drives other types of noise, such as noise from vegetation moving in the wind. Even without wind, there may be noise from branches creaking and breaking in the heat or noise from rustling leaves in the understory as animals walk through. Wind also drives waves; surf noise or noise from breaking waves is typical for coastal areas. Even without wind, moving water, such as waterfalls, can be noisy. Precipitation (i.e., rain, hail, thunder, and lightning) creates noise. Geological events such as earthquakes, seismic rumblings, and volcanic eruptions contribute noise to the terrestrial soundscape. In polar regions, melting ice and calving glaciers contribute to ambient noise.

Biotic ambient noise comes from animals in the environment. These can be of the same or different species from the target species. Several taxa call in large numbers at certain times of day and season, significantly raising ambient noise levels (e.g., chorusing cicadas, katydids, or frogs). Biologists typically think of soniferous animals as calling with specialized anatomies for sound production (i.e., syringes in birds and vocal cords in mammals). However, most animals also can produce mechanical sounds using external anatomies, such as wing-stridulation by a locust, abdomen vibration by a spider, beak-pecking by a woodpecker, teeth-chattering by a squirrel, foot-thumping by a rabbit, etc. In addition, animals can produce unintentional sounds, such as noise associated with rustling leaves as an animal walks through a forest, respiration noise, flight noise, feeding sounds, etc., not intended for communication with a conspecific. Example spectrograms for many of these sounds are found in Chap. 7 on soundscapes as well as Chap. 8 on detecting and classifying animal sounds.

Anthropogenic ambient noise is due to aircraft, road traffic, trains, ships, military activities, construction activities, etc. Increasing encroachment of human activities on animal habitats results in increased noise exposure for all taxa of animals (see Chap. 13 on noise impacts).

Ambient noise varies with time on scales of hours, days, lunar phase, season, and year. The reason is a combination of sound propagation effects and source behavior. The time of day and season of year affect sound propagation. As

explained above, sounds can be heard from farther away during the night; for example, a train can be heard in the distance at night, but not during the day. Walking in the woods during the winter, the listener can hear sounds over much greater distances than during the summer with thick vegetation. In many animals, sound-production rates are highest during the breeding season. Chorusing insects, amphibians, and birds precisely time the commencement of their cacophonies to a breeding season each year. Amphibians stop calling when they go into winter hibernation, so chorusing can stop abruptly in late autumn. Some birds migrate, so their songs are missing from the winter soundscape. Many migrating birds are soniferous and their flight calls can temporarily dominate the soundscape as they pass through an area during a spring migration (e.g., a honking flock of migrating geese or a chirping flock of starlings). Yet, other species of birds remain in temperate areas over winter and produce sounds all year long (e.g., cardinals, sparrows, and snow juncos). Tropical insects, frogs, and birds can reproduce multiple times per year, they do not migrate or hibernate, and so are soniferous throughout the year. Diurnal cycles exist in all animals with birds calling in the morning, insects in the afternoon, frogs in the evening, and nocturnal animals in the middle of the night.

### 5.3.3 The Receiver

The same factors that can affect the sender also could affect the receiver's ability to detect and interpret a signal (i.e., species, population, individual traits, age, sex, context, and ambient noise). On the species level, different species typically hear sound at different frequencies and levels. In other words, audiograms are species-specific (Fig. 5.19). Fortunately, data on hearing abilities of invertebrates, insects, reptiles, amphibians, fish, birds, and mammals continue to accumulate (see Volume 2). Nonetheless, there is some intra-species and individual variability in hearing (see Chap. 10).

In American mink (*Neovison vison*), for instance, hearing-sensitivity and frequency range changed markedly with postnatal age. Pups up to 32 days old were almost deaf, whereas three weeks later, their audiogram started to resemble that of an adult (in shape), but they remained less sensitive than adults, especially below 10 kHz (Brandt et al. 2013). There might be good reasons why hearing in young is immature. For example, a male fruit fly (*Drosophila melanogaster*) cannot hear the female's flight tone until he is physically mature enough to mate (Eberl and Kernan 2011). This ensures the female fruit fly that any pursuing male is mature. Hearing capabilities further change over an adult's life. Natural deterioration with age due to anatomical and physiological aging is a process called presbycusis. Hearing loss can also be caused by acute noise exposure at strong levels and chronic exposure to moderate noise (see Chap. 13). Hearing loss likely affects the ability of a receiver to hear and interpret a sender's message. For example, a hearing-impaired moth, which typically avoids a bat predator through an evasive flight pattern, will be easier to capture if the bat's echolocation signals are not heard.

The receiver's sex rarely influences its hearing capabilities; however, Narins and Capranica (1976, 1980) provided an example of sex differences in the auditory reception system of a Puerto Rican treefrog, the coquina frog (*Eleutherodactylus coqui*). Male and female treefrogs responded to different notes of the male's two-note, co-qui call. Females were attracted to the qui-part of the call. Males paid most attention to the co-part of the call, which was important in male–male aggressive interactions. The authors found that the inner ear basilar papilla was tuned differently in males and females; males had fewer fibers tuned to the qui-part of the call and females had fewer fibers tuned to the co-part of the call. These differences also occurred in higher-order neurons in the brain, where response decisions take place. Later studies (Mason et al. 2003) showed similar sexual differences in the middle ear of bullfrogs (*Lithobates catesbeianus*).

Ambient noise is a ubiquitous factor influencing signal reception and interpretation.

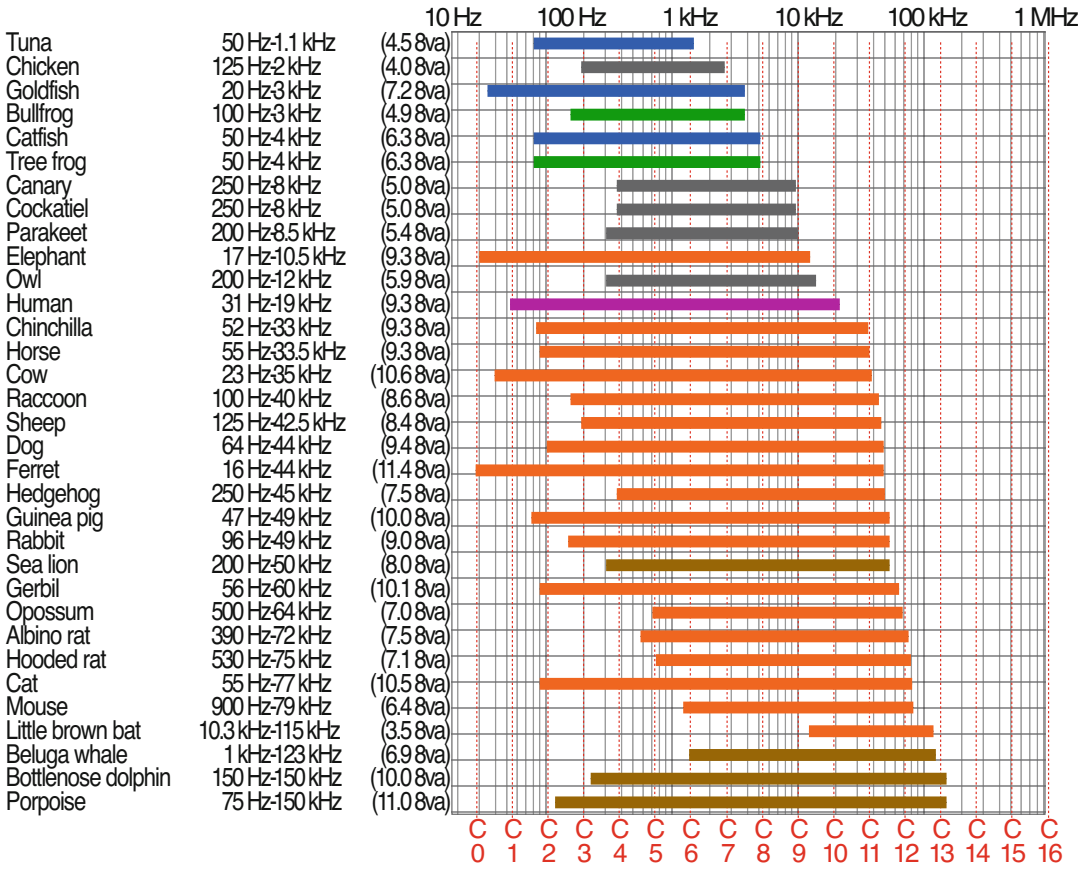



**Fig. 5.19** Hearing ranges of some animals and humans. Bars represent the approximate hearing frequency range, ordered after increasing upper frequency cut-off; blue: fish, gray: bird, green: frog, orange: terrestrial mammal, violet: human, and brown: marine mammal. The red vertical lines are the frequencies of musical notes $C_0$–$C_{16}$, for comparison. There is one octave between successive C-notes. Middle-C on a piano is $C_4$. A full-sized piano will only range from just under $C_1$ to $C_8$, with tones $>C_{11}$ being ultrasound. Data from Fay (1988), Fay and Popper (1994), Heffner (1983), Heffner and Heffner (2007), Lipman and Grassi (1942), Warfield (1973), and West (1985), previously compiled by Vanderbilt University and Louisiana State University (http://lsu.edu/deafness/HearingRange.html; accessed 6 January 2021), and plotted by Wikimedia Commons author Cmglee. https://commons.wikimedia.org/wiki/File:Animal_hearing_frequency_range.svg. Figure licensed under the Creative Commons Attribution-Share Alike 3.0 Unported license; https://creativecommons.org/licenses/by-sa/3.0/deed.en

Having experienced various forms of attenuation along its path, a signal will be audible if its amplitude remains above the power spectral density level of the ambient noise plus the critical ratio of the receiver. The critical ratio is essentially a minimum SNR needed for signal detection (see Chap. 10 for more information on the critical ratio). An even higher SNR is needed for signal discrimination, recognition, and finally, comfortable communication (Fig. 5.20; Lohr et al. 2003; Dooling et al. 2009; Dooling and Blumenrath 2013; Dooling and Leek 2018). Some birds take advantage of these limitations by producing both high-amplitude broadcast sounds and low-amplitude soft sounds. The former become public since they cover a large active space with many potential receivers whereas the latter become private as they cover a very small active space with only few receivers (Larsen 2020).

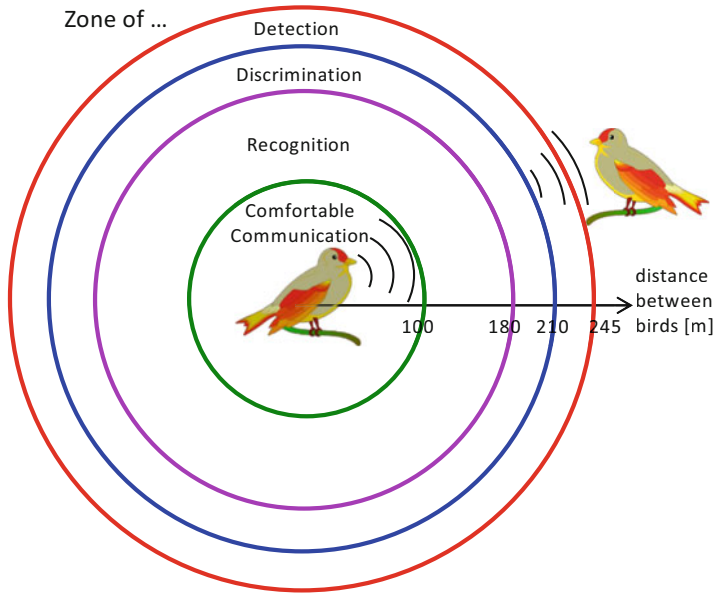



**Fig. 5.20** Sketch of the radii about a calling bird over which a broadcast public call might be detected, discriminated, and recognized. Detection (i.e., signal presence/absence) is possible over the longest ranges (i.e., lowest SNR). A higher SNR is needed for signal discrimination, then signal recognition, and finally, comfortable communication, yielding progressively shorter ranges. In louder ambient noise, the ranges will be even less. For animals with soft private calls or greater critical ratios, the radii will also be less (Erbe et al. 2016). © Erbe et al.; https://doi.org/10.1016/j.marpolbul.2015.12.007. Licensed under CC BY 4.0; https://creativecommons.org/licenses/by/4.0/

The auditory systems of some animals have built-in masking-release processes to reduce the impact of ambient noise. A spatial release from masking results from the directional hearing capabilities of the animal. If the signal arrives from a direction in which the receiver is more sensitive and if the noise arrives from a direction in which the receiver is less sensitive, then the reception directivity improves the SNR and the signal can be detected in higher ambient noise. A spatial release from masking has been demonstrated in several taxa including tropical crickets (*Paroecanthus podagrosus* and *Diatrypa* sp.; Schmidt and Römer 2011), gray treefrogs (Bee 2008), budgerigars (*Melopsittacus undulatus*; Dent et al. 1997), and pigmented Guinea pigs (*Cavia porcellus*; Greene et al. 2018). A comodulation masking release is possible if the noise is broadband and amplitude-modulated coherently across its frequencies. The animal might then utilize information about the noise from frequencies outside of the signal frequency to filter the noise within the frequency band of the signal. A comodulation masking release has been demonstrated in gray treefrogs (Bee and Vélez 2018), European starling (*Sturnus vulgaris*; Klump and Langemann 1995), and house mice (*Mus musculus*; Klink et al. 2010). Addionally, animals have a host of behavioral adaptations to optimize sound reception. For example, an animal may improve the SNR for sound arriving at its ears by approaching the source, tilting its head, adjusting its pinnae (in the case of mammals), or moving to another location away from a noise source (Nelson and Suthers 2004).

## 5.4 Summary

The Source-Path-Receiver Model (SPRM) is used widely in technical noise control and illustrates

the importance of exploring a signal at all points between the source and receiver and of understanding factors that affect the observations. This chapter developed the SPRM for the example of animal acoustic communication (also see Chap. 11). The influences of the sender's and receiver's species, age, sex, individual identity, and behavioral status were discussed. The receiving animal's hearing ability is a major factor for communication success.

Terminology related to sound propagation (or the path) was defined and basic concepts of outdoor sound propagation were developed, supported with simple equations. Several factors play an important role in sound propagation: distance between sender and receiver, air temperature, wind (direction and speed), obstacles along the path, and ground cover. The concepts of source level, received level, sound absorption, reflection, scattering, reverberation, diffraction, refraction, acoustic shadows, acoustic mirages, air temperature gradients, and wind speed gradients were illustrated. Two types of geometric spreading (i.e., spherical and cylindrical) were applied. Examples for ray tracing were provided. Ambient noise (including its abiotic, biotic, and anthropogenic sources) in terrestrial environments and its influence on both sender and receiver was discussed.

The SPRM may be applied to many other bioacoustic scenarios or studies such as animal biosonar (where the sender and receiver are the same individual; see Chap. 12) or the effects of noise on animals (where the source might be a highway; see Chap. 13). It would also be useful to consider passive acoustic monitoring (of animals or soundscapes) within the framework of the SPRM to understand the sound sources recorded, the way the environment affects the recorded soundscape, and the effects (and potential artifacts) of the recording system (i.e., the receiver; see Chaps. 2 and 7). The SPRM might also guide the bioacoustician in setting up audiometric experiments (where the source is an engineered signal; see Chap. 10). The SPRM is a fundamental concept helpful in bioacoustic study design and interpretation.

## 5.5 Additional Resources

The following sites were last accessed 3 February 2021.

- *NoiseModelling* is a free software package developed by the French Government's Centre National de la Recherche Scientifique and the Université Gustave Eiffel to produce sound maps: https://noise-planet.org/noisemodelling.html
- Dan Russell's Acoustics and Vibration Animations: https://www.acs.psu.edu/drussell/demos.html

**Acknowledgement** We wish to thank Prof. Keith Attenborough for his constructive review of this chapter.

## References

Arnold B, Wilkinson G (2011) Individual specific contact calls of pallid bats *Antrozous pallidus* attract conspecifics. Behav Ecol Sociobiol 65:1581–1593. https://doi.org/10.1007/s00265-011-1168-4

Attenborough K (2007) Sound propagation in the atmosphere. In: Rossing TD (ed) Springer handbook of acoustics. Springer, New York, pp 113–147

Attenborough K, Taherzadeh S, Bass HE, Di X, Raspet R, Becker GR, Güdesen A, Chrestman A, Daigle GA, L'Espérance A, Gabillet Y, Gilbert KE, Li YL, White MJ, Naz P, Noble JM, van Hoof HAJM (1995) Benchmark cases for outdoor sound propagation models. J Acoust Soc Am 97(1):173–191. https://doi.org/10.1121/1.412302

Attenborough K, Li KM, Horoshenkov K (2007) Predicting outdoor sound. Taylor & Francis, London

Aubin T, Jouventin P (2002) How to vocally identify kin in a crowd: the penguin model. Adv Study Behav 31:243–277

Aylor D (1972a) Noise reduction by vegetation and ground. J Acoust Soc Am 51(2):197–205

Aylor D (1972b) Sound transmission through vegetation in relation to leaf area density, leaf width, and breadth of canopy. J Acoust Soc Am 51(1):411–414

Balcombe JP, Fenton MB (1988) The communication role of echolocation calls in vespertilionid bats. In: Nachtigall PE, Moore PWB (eds) Animal Sonar, NATO ASI science (Series A: Life Sciences), vol 156. Springer, Boston, MA

Bee MA (2008) Finding a mate at a cocktail party: spatial release from masking improves acoustic mate recognition in grey treefrogs. Anim Behav 75(5):1781–1791. https://doi.org/10.1016/j.anbehav.2007.10.032

Bee MA, Vélez A (2018) Masking release in temporally fluctuating noise depends on comodulation and overall level in Cope's gray treefrog. J Acoust Soc Am 144(4): 2354–2362. https://doi.org/10.1121/1.5064362

Bradbury JW, Vehrencamp SL (2011) Principles of animal communication, 2nd edn. Sinauer Associates, Sunderland, MA

Brandt C, Malmkvist J, Nielsen RL, Brande-Lavridsen N, Surlykke A (2013) Development of vocalization and hearing in American mink (*Neovison vison*). J Exp Biol 216:3542–3550

Brumm H (2004) The impact of environmental noise on song amplitude in a territorial bird. J Anim Ecol 73: 434–440

Brumm H, Todt D (2003) Facing the rival: directional singing behaviour in nightingales. Behaviour 140(1): 43–53

Brumm H, Zollinger SA (2011) The evolution of the Lombard effect: 100 years of psychoacoustic research. Behaviour 148(11–13):1173–1198. https://doi.org/10.1163/000579511X605759

Chappuis C (1971) Un exemple de l'influence du milieu sur les émissions vocales des oiseaux: L'évolution des chants en fôret équatoriale. Terre Vie 118:183–202

Charlton BD, Ellis WA, McKinnon AJ, Cowin GJ, Brumm J, Nilsson K, Fitch WT (2011) Cues to body size in the formant spacing of male koala (*Phascolarctos cinereus*) bellows: honesty in an exaggerated trait. J Exp Biol 214:3414–3422

Dabelsteen T (1981) The sound pressure level in the dawn song of the blackbird *Turdus merula* and a method for adjusting the level in experimental song to the level in natural song. Z Tierpsychol 56(2):137–149

Dabelsteen T, Larsen ON, Pedersen SB (1993) Habitat-induced degradation of sound signals: Quantifying the effects of communication sounds and bird location on blur ratio, excess attenuation, and signal-to-noise ratio in blackbird song. J Acoust Soc Am 93(4):2206–2220

Daigle GA, Piercy JE, Embleton T (1983) Line-of-sight propagation through atmospheric turbulence near the ground. J Acoust Soc Am 74(5):1505–1513

Defrance J, Barriere N, Premat E (2002) Forest as a meteorological screen for traffic noise. Proc. 9th ICSV, Orlando, FL, USA.

Dent ML, Larsen ON, Dooling RJ (1997) Free-field binaural unmasking in budgerigars *(Melopsittacus undulatus)*. Behav Neurosci 111:590–598

Dooling RJ, Blumenrath SH (2013) Avian sound perception in noise. In: Brumm H (ed) Animal communication and noise, Animal signals and communication, vol 2. Springer-Verlag, Heidelberg, pp 229–250

Dooling RJ, Leek MR (2018) Communication masking by man-made noise. In: Slabbekoorn H, Dooling RJ, Popper AN, Fay RR (eds) Effects of anthropogenic noise on animals, Springer handbook of auditory research, vol 66, New York, pp 23–46

Dooling RJ, West EW, Leek MR (2009) Conceptual and computational models of the effects of anthropogenic noise on birds. Proc Inst Acoust 31(1):1

Eberl DF, Kernan MJ (2011) Recording sound-evoked potentials from the *Drosophila* antennal nerve. Cold Spring Harb Protoc 2011:prot5576

Eliades SJ, Wang X (2012) Neural correlates of the Lombard effect in primate auditory cortex. J Neurosci 32(31):10737–10748. https://doi.org/10.1523/JNEUROSCI.3448-11.2012

Erbe C, Reichmuth C, Cunningham KC, Lucke K, Dooling RJ (2016) Communication masking in marine mammals: a review and research strategy. Mar Pollut Bull 103:15–38. https://doi.org/10.1016/j.marpolbul.2015.12.007

Fay RR (1988) Hearing in vertebrates: a psychophysics databook. Hill-Fay Associates, Winnetka IL

Fay RR, Popper AN (1994) Comparative hearing: mammals. Springer handbook of auditory research series. Springer-Verlag, New York

Feng AS, Arch VS, Yu Z, Yu X-J, Xu Z-M, Shen J-X (2009) Neighbor–stranger discrimination in concave-eared torrent frogs, *Odorrana tormota*. Ethology 115(9):851–856

Fenton MB, Bell G (1981) Recognition of species of insectivorous bats by their echolocation calls. J Mammal 62:233–243

Fletcher NH (2004) A simple frequency-scaling rule for animal communication. J Acoust Soc Am 115:2334–2338

Galeotti P, Pavan G (1991) Individual recognition of male tawny owls (*Strix aluco*) using spectrograms of their territorial calls. Ethol Ecol Evol 3:113–126

Galeotti P, Appleby BM, Redpath SM (1996) Macro and microgeographical variations in the "hoot" of Italian and English Tawny Owls (*Strix aluco*). Ital J Zool 63: 57–64

Gannon W, Racz GR (2006) Character displacement and ecomorphological analysis of two long-eared *Myotis* (*M. auriculus* and *M. evotis*). J Mammal 87(1): 171–179

Gannon WL, Sherwin RE, de Carvalho TN, O'Farrell MJ (2001) Pinnae and echolocation call differences between *Myotis californicus* and *M. ciliolabrum* (Chiroptera: Vespertilionidae). Acta Chiropterologica 3(1):77–91

Gannon WL, O'Farrell MJ, Corben C, Bedrick EJ (2003) Call character lexicon and analysis of field recorded bat echolocation calls. In: Thomas JA, Moss CF, Vater M (eds) Echolocation in bats and dolphins. Univ Chicago Press, Chicago, IL, pp 478–484

Garcia M, Charrier I, Rendall D, Iwaniuk AN (2012) Temporal and spectral analyses reveal individual variation in a non-vocal acoustic display: the drumming display of the ruffed grouse (*Bonasa umbellus*, L.). Ethology 118(3):292–301

Garstang M (2010) Elephant infrasounds: Long-range communication. In: Brudzynski SM (ed) Handbook of mammalian vocalization: an integrative neuroscience approach. Elsevier BV, Oxford

Garstang M, Larom D, Raspe R, Lindeque M (1995) Atmospheric controls on elephant communication. J Exp Biol 198:939–951

Geiger R (1965) The climate near the ground. Harvard University Press, Cambridge, MA

Genevois F, Bretagnolle V (1994) Male blue petrels reveal their body mass when calling. Ethol Ecol Evol 6:377–383

Gerhardt HC (1991) Female mate choice in treefrogs: static and dynamic acoustic criteria. Anim Behav 42(4):615–635

Greene NT, Anbuhl KL, Ferber AT, DeGuzman M, Allen PD, Tollin DJ (2018) Spatial hearing ability of the pigmented Guinea pig (*Cavia porcellus*): minimum audible angle and spatial release from masking in azimuth. Hear Res 365:62–76. https://doi.org/10.1016/j.heares.2018.04.011

Hage SR, Jiang T, Berquist SW, Feng J, Metzner W (2013) Lombard effect in horseshoe bats. Proc Natl Acad Sci 110(10):4063–4068. https://doi.org/10.1073/pnas.1211533110

Halfwerk W, Lea AM, Guerra MA, Page RA, Ryan MJ (2016) Vocal responses to noise reveal the presence of the Lombard effect in a frog. Behav Ecol 27(2):669–676. https://doi.org/10.1093/beheco/arv204

Hardouin LA, Reby D, Bavoux C, Burneleau G, Bretagnolle V (2007) Communication of male quality in owl hoots. Am Nat 169(4):552–562

Hardouin LA, Thompson R, Stenning M, Reby D (2014) Anatomical bases of sex- and size-related acoustic variation in herring gull alarm calls. J Avian Biol 45:157–166

Heffner HH (1983) Hearing in large and small dogs: absolute thresholds and size of the tympanic membrane. Behav Neurosci 97:310–318

Heffner HE, Heffner RS (2007) Hearing ranges of laboratory animals. J Am Assoc Lab Anim Sci 46(1):20–22

Heller EJ (2013) Why you hear what you hear. Princeton University Press, Princeton, NJ

Holland J, Dabelsteen T, Pedersen SB, Paris AL (2001) Potential ranging cues contained within the energetic pauses of transmitted wren song. Bioacoustics 12(1):3–20

Jansen DA, Cant MA, Manser MB (2012) Segmental concatenation of individual signatures and context cues in banded mongoose (*Mungos mungo*) close calls. BMC Biol *10*(1):97. https://doi.org/10.1186/1741-7007-10-97

Jensen KK, Larsen ON, Attenborough K (2008) Measurements and predictions of hooded crow (*Corvus corone cornix*) call propagation over open field habitats. J Acoust Soc Am 123(1):507–518

Klink KB, Dierker H, Beutelmann R et al (2010) Comodulation masking release determined in the mouse (*Mus musculus*) using a flanking-band paradigm. JARO 11:79–88. https://doi.org/10.1007/s10162-009-0186-7

Klump GM, Langemann U (1995) Comodulation masking release in a songbird. Hear Res 87(1):157–164. https://doi.org/10.1016/0378-5955(95)00087-K

Krokstad A, Svensson UP, Strøm S (2015) The early history of ray tracing in acoustics. In: Xiang N, Sessler G (eds) Acoustics, information, and communication. Modern acoustics and signal processing. Springer, Cham

Larom DL, Garstang M, Payne RR, Lindeque M (1997) The influence of surface atmospheric conditions on the range and area reached by animal vocalizations. J Exp Biol 200:421–431

Larsen ON (1995) Acoustic equipment and sound field calibration. In: Klump GM, Dooling RJ, Fay RR, Stebbins WC (eds) Methods in comparative psychoacoustics. Birkhäuser Verlag, Basel, pp 31–45

Larsen ON (2020) To shout or to whisper? Strategies for encoding public and private information in sound signals. In: Aubin T, Mathevon N (eds) Coding strategies in vertebrate acoustic communication, Animal signals and communication, vol 7. Springer Nature Switzerland AG, Cham, pp 11–44

Larsen ON, Dabelsteen T (1990) Directionality of blackbird vocalization. Implications for vocal communication. Ornis Scand 21:37–45

Larsen ON, Radford C (2018) Acoustic conditions affecting sound communication in air and underwater. In: Slabbekoorn H, Dooling RJ, Popper AN, Fay RR (eds) Effects of anthropogenic noise on animals. Springer handbook of acoustic research. Springer, New York, pp 109–144

Larsen ON, Wahlberg M (2017) Sound and sound sources. In: Brown CH, Riede T (eds) Comparative bioacoustics: an overview. Bentham Science Publishers, Sharjah, pp 3–62

Larsson C (2000) Weather effects on outdoor sound propagation. Int J Acoust Vib 5(1):33–36

Lipman EA, Grassi JR (1942) Comparative auditory sensitivity of man and dog. Amer J Psychol 55:84–89

Lohr B, Wright TF, Dooling RJ (2003) Detection and discrimination of natural calls in masking noise by birds: estimating the active space of a signal. Anim Behav 65:763–777

Lombard É (1911) Le signe de l'élévation de la voix. Annales des Maladies de L'Oreille et du Larynx XXXVII(2):101–109

Lovell SF, Lein MR (2004) Neighbor-stranger discrimination by song in a suboscine bird, the alder flycatcher, *Empidonax alnorum*. Behav Ecol 15(5):799–804

Marten K, Marler P (1977) Sound transmission and its significance for animal vocalization. I. Temperate habitats. Behav Ecol Sociobiol 2(3):271–290

Martínez-Sala R, Rubio C, García-Raffi LM, Sánchez-Pérez JV, Sánchez-Pérez EA, Llinares J (2006) Control of noise by trees arranged like sonic crystals. J Sound Vib 291:100–106

Mason MJ, Lin CC, Narins PM (2003) Sex differences in the middle ear of the bullfrog (*Rana catesbeiana*). Brain Behav Evol 61(2):91–101

Melendez KV, Feng AS (2010) Communication calls of little brown bats display individual-specific characteristics. J Acoust Soc Am 128:919–923

Michelsen A (1978) Sound reception in different environments. In: Ali MA (ed) Sensory ecology, NATO Adv Study Inst Ser, vol 18. Plenum Press, London, pp 345–373

Michelsen A (1992) Hearing and sound communication in small animals: evolutionary adaptations to the laws of physics. In: Webster DB, Fay RR, Popper AN (eds) The evolutionary biology of hearing. Springer-Verlag, New York, pp 61–77

Michelsen A, Larsen ON (1983) Strategies for acoustic communication in complex environments. In: Huber F, Markl H (eds) Neuroethology and behavioral physiology. Springer-Verlag, Berlin, pp 321–331

Miller EH, Williams J, Jamieson SE, Gilchrist HG, Mallory ML (2007) Allometry, bilateral asymmetry, and sexual differences in the vocal tract of common eiders *Somateria mollissima* and king eiders *S. spectabilis*. J Avian Biol 38:224–233

Mitani JC, Hasegawa T, Gros-Louis J, Marler P, Byrne R (1992) Dialects in wild chimpanzees? Am J Primatol 27:233–243. https://doi.org/10.1002/ajp.1350270402

Narins PM, Capranica RR (1976) Sexual differences in the auditory system of the tree frog *Eleutherodactylus coqui*. Science 192(4237):378–380

Narins PM, Capranica RR (1980) Neural adaptations for processing the two-note call of the Puerto Rican tree frog, *Eleutherodactylus coqui*. Brain Behav Evol 17: 48–66

Nelson BS, Suthers RA (2004) Sound localization in a small passerine bird: discrimination of azimuth as a function of head orientation and sound frequency. J Exp Biol 207:4121–4133

O'Farrell MJ, Miller BW, Gannon WL (1999) Qualitative identification of free-flying bats using the Anabat detector. J Mammal 80(1):11–23

Ottemöller L, Evers LG (2008) Seismo-acoustic analysis of the Buncefield oil depot explosion in the UK, 2005 December 11. Geophys J Int 172(3):1123–1134

Price MA, Attenborough K, Heap NW (1988) Sound attenuation through trees: measurements and models. J Acoust Soc Am 84:1836–1844

Reby D, McComb K (2003) Anatomical constraints generate honesty: acoustic cues to age and weight in the roars of red deer stags. Anim Behav 65:519–530

Riede T, Fisher JH, Goller F (2010) Sexual dimorphism of the zebra finch syrinx indicates adaptation for high fundamental frequencies in males. PLoS One 5:e11368

Römer H (1998) The sensory ecology of acoustic communication in insects. In: Hoy RR, Popper AN, Fay RR (eds) Comparative hearing: insects. Springer handbook of auditory research. Springer-Verlag, New York, pp 63–96

Ross CD (2000) Outdoor sound propagation in the US Civil War. Appl Acoust 59:137–147

Russ J, Jones G, Mackie I, Racey P (2004) Interspecific responses to distress calls in bats (Chiroptera: Vespertilionidae): A function for convergence in call design? Anim Behav 67:1005–1014. https://doi.org/10.1016/j.anbehav.2003.09.003

Schmidt AKD, Römer H (2011) Solutions to the cocktail party problem in insects: selective filters, spatial release from masking and gain control in tropical crickets. PLoS One 6(12):e28593. https://doi.org/10.1371/journal.pone.0028593

Searby A, Jouventin P, Aubin T (2004) Acoustic recognition in macaroni penguins: an original signature system. Anim Behav 67:615–625

Stoeger-Horwath AS, Stoeger S, Schwammer HM, Kratochvil H (2007) Call repertoire of infant African elephants: first insights into the early vocal ontogeny. J Acoust Soc Am 121(6):3922–3931

Suthers RA (1994) Variable asymmetry and resonance in the avian vocal tract: a structural basis for individually distinct vocalizations. J Comp Physiol A 175:457–466

ter Hofstede HM, Ratcliffe JM (2016) Evolutionary escalation: the bat-moth arms race. J Exp Biol 219(11): 1589–1602. https://doi.org/10.1242/jeb.086686

Thomas JA, Kuechle V (1982) Quantitative analysis of the underwater repertoire of the Weddell seal (*Leptonychotes weddellii*). J Acoust Soc Am 72: 1730–1738

Trefry SA, Hik DS (2010) Variation in pika (*Ochotona collaris, O. princeps*) vocalizations within and between populations. Ecography 33:784–795

Van Staaden M, Römer H (1997) Sexual signalling in bladder grasshoppers: tactical design for maximizing calling range. J Exp Biol 200:2597–2608

Vannoni E, McEligott AG (2007) Individual acoustic variation in fallow deer (*Dama dama*) common and harsh groans: a source-filter theory perspective. Ethology 113:223–234

Wahlberg M, Larsen ON (2017) Propagation of sound. In: Brown CH, Riede T (eds) Comparative bioacoustics: an overview. Bentham Science Publishers, Sharjah, pp 63–121

Warfield D (1973) The study of hearing in animals. In: Gay W (ed) Methods of animal experimentation, IV. Academic Press, London, pp 43–143

West CD (1985) The relationship of the spiral turns of the cochlea and the length of the basilar membrane to the range of audible frequencies in ground dwelling mammals. J Acoust Soc Am 77:1091–1101

Wiley RH (2009) Signal transmission in natural environments. In: Squire LR (ed) Encyclopedia of neuroscience, vol 8. Academic Press, Oxford, pp 827–832

Wiley RH, Richards DG (1978) Physical constraints on acoustic communication in the atmosphere:

Implications for the evolution of animal vocalizations. Behav Ecol Sociobiol 3:69–94

Wiley RH, Richards DG (1982) Adaptations for acoustic communication in birds: sound transmission and signal detection. In: Kroodsma DE, Miller EH, Quellet H (eds) Acoustic communication in birds. Academic Press, New York, pp 131–181

Wilkinson GS, Boughman JW (1998) Social calls coordinate foraging in greater spear-nosed bats. Anim Behav 55:337–350

Yorzinski JL, Patricelli GL (2009) Birds adjust acoustic directionality to beam their antipredator calls to predators and conspecifics. Proc R Soc B 277:923–932

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



# Introduction to Sound Propagation Under Water

6

Christine Erbe, Alec Duncan, and Kathleen J. Vigness-Raposa

## 6.1 Introduction

It is imperative that bioacousticians who work in aquatic environments have a basic understanding of sound propagation under water. Whether the topic is the function of humpback whale song, echolocation in wild bottlenose dolphins, the masking of grey whale sounds by ship noise, the role of chorusing in fish spawning behavior, the effects of seismic surveying on benthic organisms, or the capability of an echosounder to track a school of fish, the way in which sound propagates through the ocean affects how we can use sound to study animals, how sound we produce impacts animals, and how animals use sound.

Aquatic fauna has evolved to use sound for environmental sensing, navigation, and communication. This is because water conducts sound very well (i.e., fast and far), while light propagates poorly under water. Visual sensing based on sun- or moonlight is limited to the upper few meters of water. And while water transports chemicals, chemoreception is most effective over short ranges, where chemical concentration is high. Furthermore, sound can be detected from all directions, providing omnidirectional alerting of activities happening in the environment.

Given that sound may propagate over very long ranges with little loss, a myriad of sounds is commonly heard at any one place. These sounds may be grouped by origin: abiotic, biotic, and anthropogenic. Natural, geophysical, abiotic sound sources include wind blowing over the ocean surface, rain falling onto the ocean surface, waves breaking on the beach, polar ice breaking under pressure and temperature influences, subsea volcanoes erupting, subsea earthquakes rumbling along the seafloor, etc. Biotic sound sources include singing whales, chorusing fishes, feeding urchins, and crackling crustaceans. Anthropogenic sources of sound include ships, boats, fish-finding echosounders, oil rigs, gas wells, subsea mines, dredgers, trenchers, pile drivers, naval sonar, seismic surveys, underwater explosions, etc.

As these sounds travel from their source through the environment, they may follow multiple propagation paths. Sounds may be reflected at the sea surface and seafloor. Some sound may travel through the seafloor and radiate back into the water some distance away. Sound is scattered by scatterers in the water (such as gas bubbles or fish swim bladders). Sound bends as the ocean is layered with pressure, temperature, and salinity changing as a function of depth, and with freshwater inputs. All of these phenomena depend on the frequency of sound. The spectrum of broadband sound changes, too, as acoustic energy at

C. Erbe (✉) · A. Duncan
Centre for Marine Science and Technology, Curtin University, Perth, WA, Australia
e-mail: c.erbe@curtin.edu.au; A.J.Duncan@curtin.edu.au

K. J. Vigness-Raposa
INSPIRE Environmental, Newport, RI, USA
e-mail: kathy@INSPIREenvironmental.com

© The Author(s) 2022

C. Erbe, J. A. Thomas (eds.), *Exploring Animal Behavior Through Sound: Volume 1*,
https://doi.org/10.1007/978-3-030-97540-1_6

high frequencies is more readily scattered and absorbed than energy at low frequencies. The receiver of sound can thus infer information not just about the source of sound but also about the environment's complexity.

Understanding the physics of sound in water is an important step in studies of aquatic animal sound usage and perception, whether these are conspecific social sounds, predator sounds, prey sounds, navigational clues, environmental sounds, or anthropogenic sounds. It is also critical for the study of impacts of sound on aquatic fauna, and for using passive or active acoustic tools for monitoring aquatic fauna and mapping biodiversity. The goal of this chapter is to introduce the basic concepts of sound propagation under water.

## 6.2 The Sonar Equation

The sonar equation was developed by the US Navy to assess the performance of naval sonar systems. These sonar systems were designed to detect foreign submarines. The sonar emits an acoustic signal under water and listens to returning echoes. The time of arrival and acoustic features of the echo may determine not only from what target the signal reflected, but also the range and speed of the target. The term "sonar" stands for "SOund Navigation And Ranging."

There are numerous forms of the sonar equation. What they all have in common is that (1) they each represent an equation of energy conservation, meaning that the total acoustic energy on either side of the equation is the same; and (2) all of the terms in the equation are expressed in decibel (dB). The sonar equation with its original terms as defined in Urick (1983) allows an easy conceptual exploration of various scenarios encountered in bioacoustics. The definitions and notations of some of the terms are more mathematically specific in the recent underwater acoustics terminology standard (ISO 18405)[1].

### 6.2.1 Propagation Loss Form

As sound propagates through the ocean, it loses energy, termed propagation loss (PL[2]). A simple form of the sonar equation equates PL to the difference between the source level (SL) and the received level (RL) of sound (Urick 1983):

$$PL = SL - RL \text{ (propagation loss form)} \quad (6.1)$$

*SL* was defined by Urick as $10\log_{10}$ of the ratio of source intensity to reference intensity (see Chap. 4). *RL* was equal to $10\log_{10}$ of the ratio of received intensity to reference intensity. *PL* was computed as $10\log_{10}$ of the ratio of source intensity to received intensity.

For example, a whale-watching boat might have $SL = 160$ dB re 1 μPa$^2$ (in terms of mean-square pressure, which is proportional to intensity; see Chap. 4) and be located 100 m from a group of whales. If *PL* in this environment and over this range is 40 dB, then *RL* at the whales is 120 dB re 1 μPa$^2$ (Erbe 2002; Erbe et al. 2016a).

### 6.2.2 Signal-to-Noise Ratio Form

Another simple form of the sonar equation relates the RL of a signal to the background noise level ($NL = 10\log_{10}$ of the ratio of noise intensity to reference intensity):

$$SNR = RL - NL \text{ (signal-to-noise ratio form)} \quad (6.2)$$

*SNR* is the level of the signal-to-noise ratio, expressed in dB. For example, a call from a whale might have a received level $RL = 105$ dB re 1 μPa$^2$ at another whale; however, background noise at the time might be $NL = 115$ dB re 1 μPa$^2$ over the frequency band of the call. The *SNR* is −10 dB. Can the whale still hear the other one or does the noise mask the call?

Because the *SNR* is a negative number in this example, if one was just considering the relative levels of signal and noise, the animals would not

[1] International Organization for Standardization. (2017). *Underwater acoustics—Terminology (ISO 18405)*. Geneva, Switzerland.

[2] In this chapter, we italicize variables, but keep abbreviations as regular font; so PL is an abbreviation while *PL* is a variable.

be able to hear one another because the background noise level is much greater than the received signal level. However, animals (and sonar systems) can take advantage of spectral and temporal characteristics of a received sound, as is explained below. Therefore, in the example of beluga whales (*Delphinapterus leucas*) trying to communicate in icebreaker noise, the listening whale can indeed detect the call, because of the different spectral and temporal structures of call and noise (Erbe and Farmer 1998).

### 6.2.3 Forms to Assess Communication Masking

Acoustic communication under water remains an area of active research. In the conceptual model of Fig. 6.1, one animal (the sender) emits a signal, which travels through the habitat to the location of the receiver. Whether the receiver can hear the message depends on a number of factors that relate to the sender, the habitat, and the receiver. The level and spectral features of the signal will affect how far it propagates and how well it can be detected above the ambient noise in the environment. The locations of sender and receiver matter, not just the range between the two animals, but also at which depth each happens to be located. If the two animals are oriented towards each other, directional emission and reception capabilities will enhance signal detection. The environment changes the level and spectral characteristics of the signal by reflection, refraction, scattering, absorption, and spreading losses. The detection capabilities of the receiver can be quantified by the detection threshold, critical ratio, and other factors. Ambient noise in the environment can initiate anti-masking strategies at both the sender (e.g., increasing the source level) and receiver (e.g., orienting towards the signal). A sonar equation can be constructed to investigate each of these factors, as outlined in the following sections.

The basic sonar relation for the communication scenario in Fig. 6.1 is:

$$SL - PL - NL > DT \text{ (basic signal detection form)},$$

where $DT$ is the detection threshold of the receiver, expressed in dB. A sound is deemed detectable if the expression on the left side exceeds the detection threshold. In the absence of noise, $DT$ equals the audiogram. Audiograms are measured by exposing an animal to pure-tone

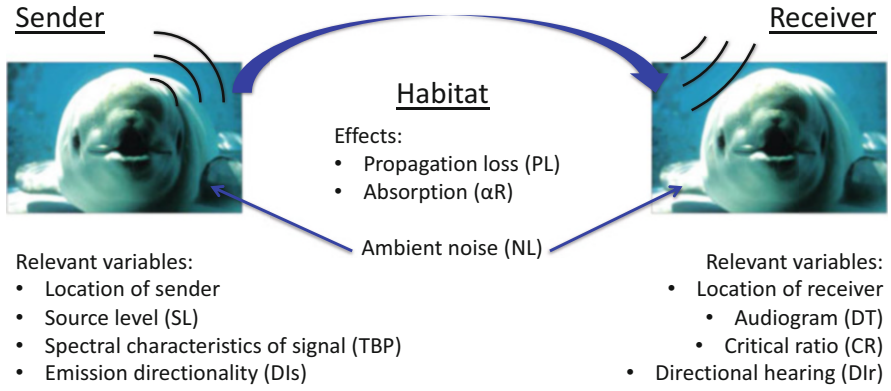



**Fig. 6.1** Sketch of the factors related to acoustic communication in natural (not just aquatic) environments and their corresponding terms in the sonar equation: source level (SL), time-bandwidth product (TBP), sender directivity index (DIs), propagation loss (PL), absorption (absorption coefficient α multiplied by range R), noise level (NL), and receiver detection threshold (DT), critical ratio (CR), and directivity index (DIr). Modified from Erbe et al. (2016c); © Erbe et al. (2016); https://www.sciencedirect.com/science/article/pii/S0025326X15302125. Published under CC BY 4.0; https://creativecommons.org/licenses/by/4.0/

signals of varying levels. The RL that is just detectable defines the audiogram at that frequency (see Chap. 10 for a more thorough definition of audiogram):

$$RL = DT \text{ (audiogram form)}$$

The mammalian auditory system acts as a bank of overlapping bandpass filters and the listener focuses on the auditory band that receives the highest SNR (Moore 2013). Under the equal-power assumption (Fletcher 1940), a signal is detected if its power is greater than the noise power in any of the auditory bands. So, for any auditory band,

$$RL - NL > 0 \text{ (within an auditory band)} \quad (6.3)$$

Communication signals of many species, including birds and marine mammals (Erbe et al. 2017a), are commonly tonal, while noise is commonly broadband. In order to assess the risk of communication masking, the critical ratio (CR) is a useful quantity that has been measured in humans and animals. The CR is the level difference between the mean-square sound pressure level (SPL) of a tone and the mean-square sound pressure spectral density level of broadband noise when the tone is just audible (American National Standards Institute 2015). Conceptually, the CR quantifies the ability of the auditory system to focus on a narrowband (tonal) signal. It captures how many of the noise frequencies surrounding the tone frequency are effective at masking the tone, and the resulting band of frequencies has been termed the Fletcher critical band (American National Standards Institute 2015). A narrowband signal is thus detectable, if

$$RL - CR > NL_f \text{ (critical ratio form)} \quad (6.4)$$

$RL$ is the tone level in dB re 1 μPa$^2$, $NL_f$ is the noise mean-square pressure spectral density level in dB re 1 μPa$^2$/Hz, and $CR$ is measured in dB re 1 Hz (see p. 29 in Erbe et al. 2016c).

In the above-mentioned study with beluga whales communicating amidst icebreaker noise, the beluga whale call consisted of a sequence of six tones with overtones from 800 to 1800 Hz, and the icebreaker's bubbler system noise was broadband and relatively unstructured in frequency and time (Fig. 6.2) (Erbe and Farmer 1998). The bandwidth of the call, expressed in dB, was $10\log_{10}(1800–800) = 30$ dB re 1 Hz (see Chap. 4 for definitions and formulae). Given

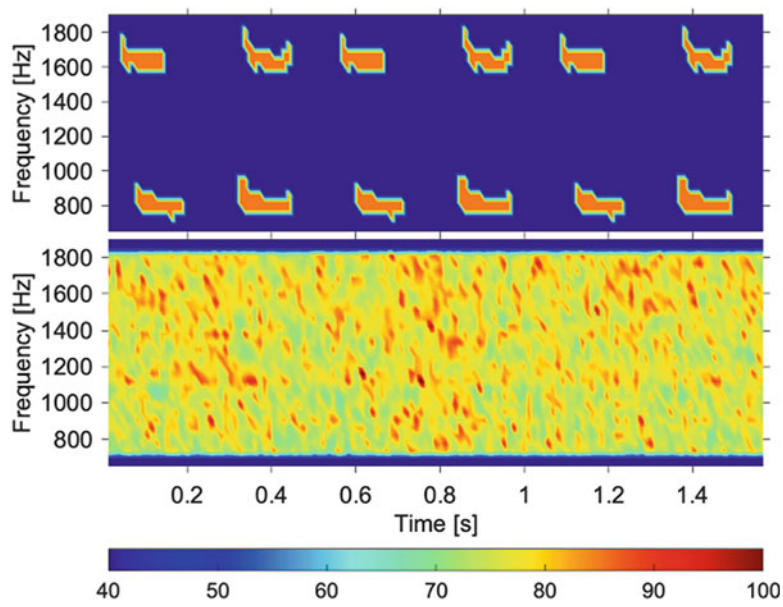



**Fig. 6.2** Spectrograms of the lower two harmonics of a beluga whale call (top panel) and an icebreaker's bubbler system noise (bottom panel). Colorbar in dB re 1 μPa$^2$/Hz. The broadband levels are $RL = 105$ dB re 1 μPa$^2$ for the call and $NL = 115$ dB re 1 μPa$^2$ for the noise

$NL = 115$ dB re 1 μPa$^2$ over the bandwidth of the call, $NL_f$ was equal to $NL$ (115 dB re 1 μPa$^2$) minus the bandwidth (30 dB re 1 Hz): $NL_f = 85$ dB re 1 μPa$^2$/Hz. Beluga whales have a CR of approximately 15 dB re 1 Hz at 800 Hz, therefore, the call with $RL = 105$ dB re 1 μPa$^2$ was audible, because Eq. (6.4) was satisfied (Erbe 2008; Erbe and Farmer 1998): 105–15 > 85.

In studies on critical ratios and in the beluga whale experiments (Erbe and Farmer 1998; Erbe 2000), signal and noise were broadcast by the same loudspeaker and thus arrived at the listener from the same direction. If the caller and the noise are spatially separated, then there is an additional processing gain in the sonar equation: the receiver's directivity index *DIr*:

$$RL - CR + DIr - NL_f > 0$$
(critical ratio form with directivity index)

The *DIr* is defined as $10\log_{10}$ of the ratio of the intensity measured by an omnidirectional receiver to that of a directional receiver. Directivity indices increase with frequency and values up to 19 dB have been measured for communication sounds in marine mammals. The associated spatial release from masking should be considered in environmental impact assessments of underwater noise (Erbe 2015). Directivity indices are even greater at higher frequencies used by dolphins during echolocation (Fig. 6.3).

### 6.2.4 Form for Biomass Surveying

Surveys for animals ranging from zooplankton to fish and sharks may use an echosounder, fish finder, or sonar (e.g., Parsons et al. 2014; Kloser et al. 2013). In this scenario, the echosounder emits a signal, which travels to the fish, where some of it is reflected. How much of the signal is reflected is expressed by the target strength (TS), defined as $10\log_{10}$ of the ratio of echo intensity to incident intensity (Urick 1983). The reflected signal travels to the receiver, which has a specific *DT* and *DIr*. The receiver is typically co-located with the source, so that the signal travels the same path twice and thus experiences twice the PL. The fish is detected if the following sonar equation is satisfied:

$$SL - 2\,PL + TS - NL > DT - DIr$$
(two − way sonar surveying form)

Target strength will vary for each type of animal, as well as with the number of animals in the group and their orientation relative to the echosounder. Figure 6.4 shows reflected signals received on a REMUS autonomous underwater vehicle. Individual animals are observed in two aggregations, with two dolphins swimming within one of the aggregations. Researchers are using cameras on the same platforms to better understand the information contained in reflected

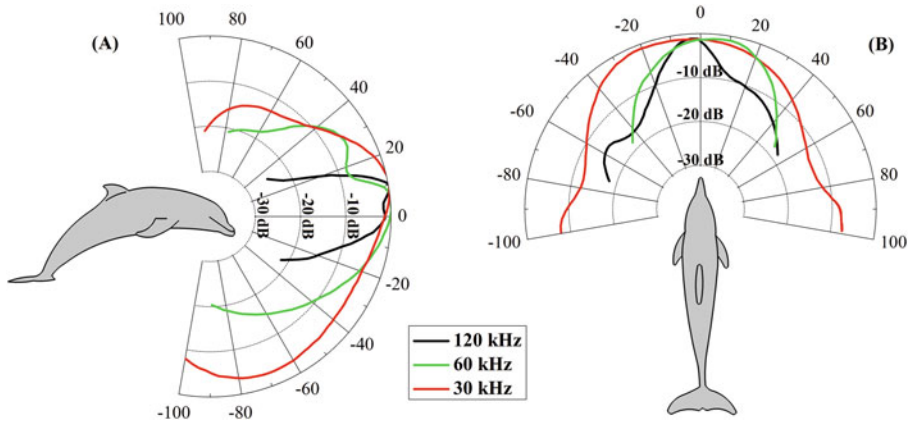



**Fig. 6.3** Sketches of the receiving directivity pattern of a bottlenose dolphin (*Tursiops truncatus*) in the vertical (**a**) and horizontal (**b**) planes. Courtesy of Chong Wei after data in (Au and Moore 1984)

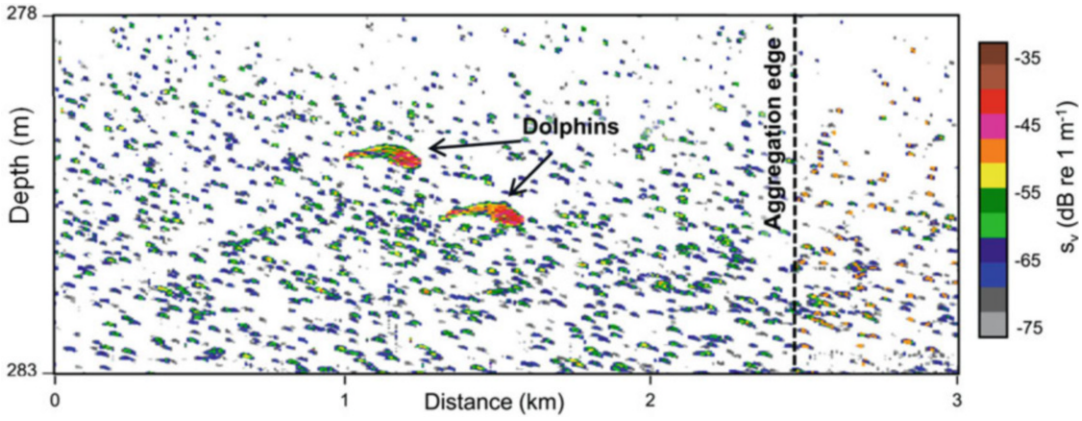



**Fig. 6.4** Echosounder image of marine fauna in two aggregations, with two dolphins being in the aggregation on the left. Colors represent acoustic target strength and the shapes of the two dolphins can easily be recognized by their high reflectivity (Benoit-Bird et al. 2017). © Benoit-Bird et al. 2017; https://aslopubs.onlinelibrary.wiley.com/doi/full/10.1002/lno.10606. Published under CC BY 4.0; https://creativecommons.org/licenses/by/4.0/

signals and ultimately convert that information into species classifications and estimates of biomass (Benoit-Bird and Waluk 2020).

## 6.3 The Layered Ocean

The speed of sound in sea water increases with increasing temperature $T$ [°C], salinity $S$ (measured in practical salinity units [psu]) and hydrostatic pressure, which in the ocean is proportional to depth $D$ [m]. The approximate change in the speed of sound $c$ [m/s] with a change in each property is:

- Temperature changes by 1 °C → $c$ changes by 4.0 m/s
- Salinity changes by 1 psu → $c$ changes by 1.4 m/s
- Depth (pressure) changes by 1 km → $c$ changes by 17 m/s

Maps of sea surface temperature and salinity for the northern hemisphere summer show considerable variation (Fig. 6.5). However, temperature and salinity vary much more rapidly with depth than they do in the horizontal plane, so the ocean can often be thought of as a stack of horizontal layers, with each layer having different properties. Vertical profiles of these quantities are therefore very useful for understanding how sound will propagate in different geographical regions.

### 6.3.1 Temperature and Salinity Profiles

In non-polar regions (red curves in Fig. 6.6), the main source of heat entering the ocean is solar. The sun heats the near-surface water, making it less dense and suppressing convection. A surface mixed layer with nearly constant temperature and salinity is formed by mechanical mixing due to surface waves and is typically 20–100 m thick. Below that, the temperature drops rapidly in a region known as the thermocline, before becoming almost constant at a temperature of about 2 °C in the deep isothermal layer that extends from a depth of about 1000 m to the ocean floor.

Seasonal changes in solar radiation together with the ocean's considerable thermal lag (due to its great heat capacity) can complicate this simple picture, but most of these changes only affect the top few hundred meters of the water column, changing the detailed structure of the mixed layer and the upper part of the thermocline.

In polar regions (blue curves in Fig. 6.6), the situation is quite different. There is a net loss of

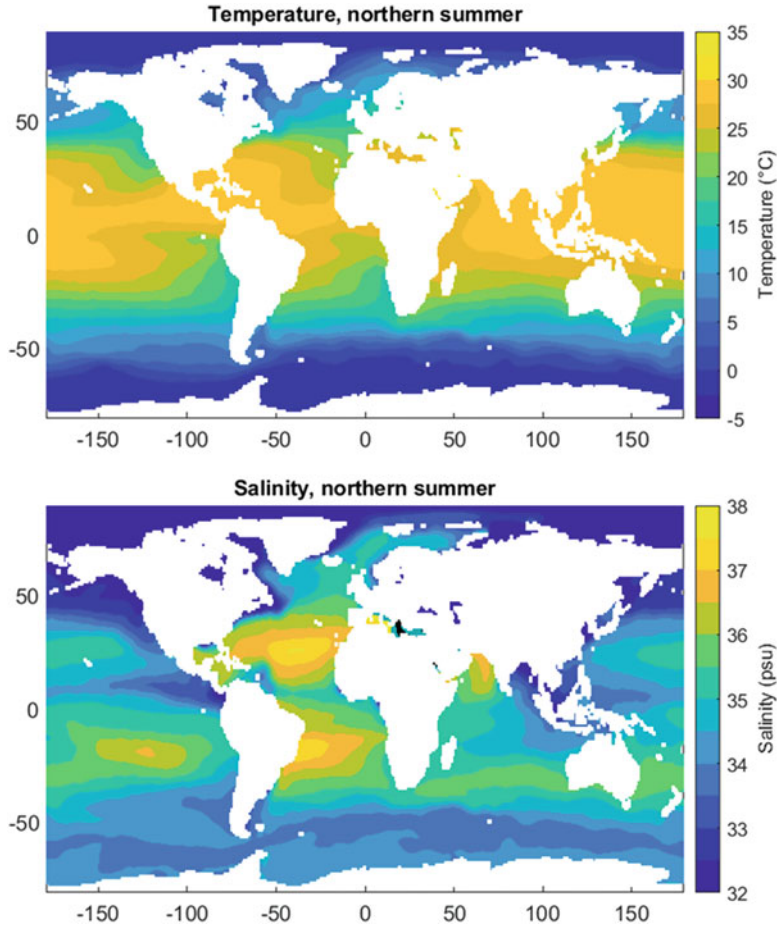



**Fig. 6.5** Maps of sea surface temperature (top) and salinity (bottom) for the northern hemisphere summer, averaged over the period 2005 to 2017. Data were taken from the *World Ocean Atlas* (Locarnini et al. 2018; Zweng et al. 2018)

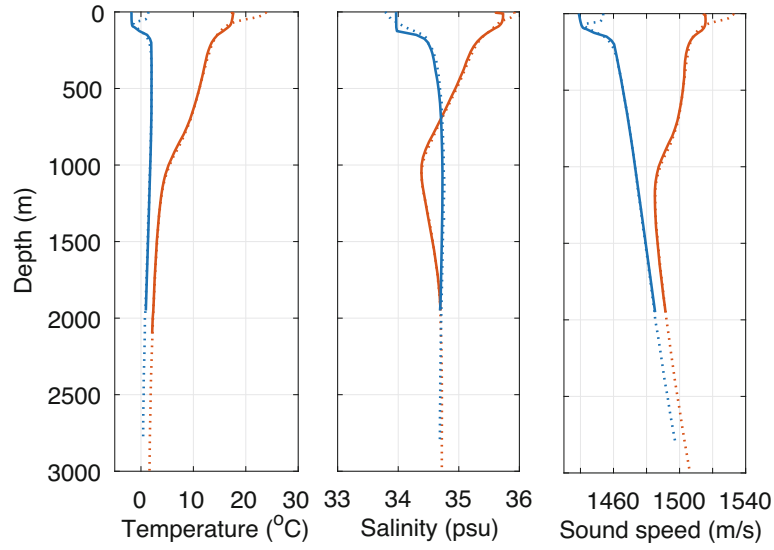



**Fig. 6.6** Depth profiles of temperature, salinity, and sound speed from the open ocean based on the *World Ocean Atlas* (Locarnini et al. 2018; Zweng et al. 2018) seasonal decadal average data for the austral winter (solid) and austral summer (dotted). Red curves are for 30.5°S, 74.5°E and are representative of non-polar ocean profiles. Blue curves are for 60.5°S, 74.5°E and are representative of polar ocean profiles

heat from the sea surface, which results in a temperature profile in the upper part of the ocean that increases with increasing depth from a minimum of about −2 °C at or (in summer) slightly below the surface.

Salinity typically changes by only a small amount with depth, and in most parts of the ocean is between 34 and 36 psu. As a result, the sound speed is usually determined by temperature and depth, however, salinity can have an important effect on sound speed in situations where it changes abruptly. Examples include locations where there is a large freshwater outflow into the ocean from a river, or in estuaries where it is common to have a wedge of dense, saline water underlying a surface layer of freshwater. In polar regions, the salinity of near-surface water can vary considerably depending on whether sea ice is forming, a process that excludes salt and therefore increases salinity in the water below the ice. When sea ice melts, freshwater is released, reducing near-surface salinity.

### 6.3.2 Sound Speed Profiles

The following equation is one of a number of equations of varying complexity that can be found in the literature relating the speed of sound to temperature, salinity, and depth (Mackenzie 1981). It is valid for temperatures from −2 to 30 °C, salinities of 30 to 40 psu, and depths from 0 to 8000 m.

$$\begin{aligned} c = {} & 1448.96 + 4.591\,T - 5.304 \times 10^{-2}\,T^2 \\ & + 2.374 \times 10^{-4}\,T^3 + 1.340\,(S-35) \\ & + 1.630 \times 10^{-2}\,D + 1.675 \times 10^{-7}\,D^2 \\ & - 1.025 \times 10^{-2} T(S-35) - 7.139 \\ & \times 10^{-13}\,TD^3\ [\mathrm{m/s}] \end{aligned}$$

Sound speed profiles computed from the typical temperature and salinity profiles are also plotted in Fig. 6.6.

In non-polar waters, the sound speed may increase slightly with depth in the mixed layer due to its pressure dependence, however, diurnal heating and cooling effects can eliminate or enhance this effect. As explained later in this chapter, whether or not there is a distinct increase in sound speed with depth in the mixed layer determines whether there is a surface duct, which has a considerable impact on acoustic propagation from near-surface sound sources and to near-surface receivers.

Below the mixed layer, the rapid reduction in temperature with depth (i.e., in the thermocline) results in sound speed also reducing until, at a depth of about 1000 m, the temperature becomes nearly constant. In the deeper isothermal layer, the increasing pressure results in the sound speed starting to increase with depth. There is therefore a minimum in the sound speed in non-polar waters at a depth of approximately 1000 m, which, as will be seen later, is important for long-range sound propagation.

In polar waters, the temperature and pressure both increase with increasing depth, so the sound speed also increases, which results in a strong surface duct. However, in the Arctic Ocean, the existence of water masses with different properties entering from the Pacific and Atlantic oceans can lead to more complicated sound speed profiles.

Temperature and salinity profiles for the world's oceans can be found in the *World Ocean Atlas*[3] (Locarnini et al. 2018; Zweng et al. 2018). These are based on averages of a large amount of measured data and are very useful for calculating estimated sound speed profiles for particular locations for particular months or seasons of the year. The real ocean is, however, highly variable; particularly the upper thermocline and mixed layer, which can change on time scales of hours, and in some extreme cases, tens of minutes, so there is no substitute for in situ measurements of temperature and salinity profiles to support acoustic work.

[3] *World Ocean Atlas* https://www.nodc.noaa.gov/OC5/woa18/; accessed 30 September 2020.

## 6.4 Propagation Loss

The apparent simplicity of the propagation loss term (i.e., *PL*) in the various sonar equations hides a great deal of complexity. There are a few special situations in which *PL* can be calculated quite accurately using simple formulae, and a few more in which it might be possible to obtain a reasonable estimate using a more complicated equation, but for everything else, these simple approaches can lead to large errors, and it is necessary to resort to numerical modeling. To further complicate matters, there are a number of different types of numerical models used for propagation loss calculations, each with its own assumptions and limitations, and it is important to be familiar with these so that the most appropriate model can be used for a given task.

### 6.4.1 Geometric Spreading Loss

The most basic concept of propagation loss is that of geometric spreading, which accounts for the fact that the same sound power is spread over a larger surface area as the sound propagates further from the source. The intensity is the sound power per unit area (see Chap. 4), so the increase in surface area results in a reduction in intensity. The simplest case is when the source is small compared to the distances involved, the sound speed is constant, and the boundaries (i.e., sea surface, seabed, and anything else that might reflect sound) are sufficiently far away that reflected energy can be ignored. In this situation, the acoustic wavefront forms the surface of a sphere. As the wavefront propagates outward, the radius $r$ of the sphere increases, the surface area of the sphere increases in proportion to $r^2$, and therefore the intensity decreases inversely proportional to $r^2$. This leads to the well-known spherical spreading equation for PL:

$$PL = 20\log_{10}(r/1\text{m}) \tag{6.5}$$

Equation (6.5) is also applicable to calculating geometric spreading loss for sound radiated by a directional source, such as an echosounder transducer, or a dolphin's biosonar, providing the range is sufficiently large (i.e., the receiver is in the acoustic far-field of the source; see Chap. 4), and the above assumptions are all met.

Another situation in which spreading loss can be calculated analytically is when the sound is constrained in one dimension by reflection and/or refraction, so it can only spread in the other two dimensions. In underwater acoustics, this most commonly happens when the sound is constrained in the vertical direction by the sea surface or seafloor, but can still spread in the horizontal plane. The result is that the acoustic wavefront forms the surface of a cylinder, the area of which is proportional to the range. The intensity is therefore inversely proportional to the range, and the PL is given by the cylindrical spreading equation:

$$PL = 10\log_{10}(r/1\text{m}) \tag{6.6}$$

Some situations in which cylindrical spreading can occur are discussed later in this chapter, but it should be noted that Eq. (6.6), strictly speaking, only applies at *all* ranges from the source in the highly unusual case that the source is a vertical line source that spans the entire depth interval into which the sound is constrained, and that no sound is lost into either the upper or lower layers.

For the much more common case of a small source, the sound will undergo spherical spreading at short ranges where the boundaries have no effect, followed by cylindrical spreading at long ranges where the fact that the source has a small vertical extent is of little consequence. In between, there will be a transition region in which neither formula is accurate. This situation can be approximated by assuming a sudden transition from spherical to cylindrical spreading at a "transition range" $r_t$. Equation (6.7) applies only to ranges $r \gg r_t$ and still makes the assumption that there are no losses at the boundaries.

$$\begin{aligned} PL &= 20\log_{10}\left(\frac{r_t}{1\text{m}}\right) + 10\log_{10}\left(\frac{r}{r_t}\right) \\ &= 10\log_{10}\left(\frac{r_t}{1\text{m}}\right) + 10\log_{10}\left(\frac{r}{1\text{m}}\right) \end{aligned} \tag{6.7}$$

In shallow-water situations, some authors recommend using a transition range equal to the

water depth; however, while useful for very rough *PL* estimates, this approach should be adopted with caution as the best choice will depend on the characteristics of the seabed. The only way to accurately determine $r_t$ for a given situation is to carry out numerical propagation modeling, in which case you might as well use that to directly determine the propagation loss, removing the need for (Eq. 6.7) and its inherent inaccuracies.

### 6.4.2 Absorption Loss

When a sound wave propagates through water, it results in a periodic motion of the molecules present in the water, and the slight friction within and between them converts some of the sound energy into heat, reducing the intensity of the sound wave. This is called absorption loss and results in a propagation loss that is proportional to the range traveled:

$$PL = \alpha r_{\mathrm{km}} \tag{6.8}$$

where $r_{\mathrm{km}}$ is the range in kilometers and $\alpha$ is the absorption coefficient in dB/km. The propagation loss due to absorption must be added to the propagation loss due to geometrical spreading described in Sect. 6.4.1.

A commonly used formula for $\alpha$ is:

$$\begin{aligned}\alpha = {} & 0.106\frac{f_1 f^2}{f_1^2+f^2}e^{(pH-8)/0.56} \\ & + 0.52\left(1+\frac{T}{43}\right)\frac{S}{35}\frac{f_2 f^2}{f_2^2+f^2}e^{-z/6} \\ & + 4.9\times 10^{-4} f^2 e^{-(T/27+z/17)}\end{aligned} \tag{6.9}$$

with $f_1 = 0.78(S/35)^{1/2}\mathrm{e}^{T/26}$ and $f_2 = 42\mathrm{e}^{T/17}$; $f$ [kHz], $\alpha$[dB/km]

valid for $-6 < T < 35°\mathrm{C}$ ($S = 35$psu, pH = 8, $z = 0$)
$7.7 < \mathrm{pH} < 8.3$ ($T = 10°\mathrm{C}$, $S = 35$psu, $z = 0$)
$5 < S < 50$psu ($T = 10°\mathrm{C}$, pH = 8, $z = 0$)
$0 < z < 7$km ($T = 10°\mathrm{C}$, $S = 35$psu, pH = 8)

(François and Garrison 1982a, b; Ainslie and McColm 1998).

The absorption coefficient increases with frequency (Fig. 6.7). At low frequencies, it is dominated by molecular relaxation of two minor constituents of seawater: $B(OH)_3$ and $MgSO_4$, whereas above a few hundred kHz, it is primarily due to the water's viscosity.

In summary, Fig. 6.8 compares how propagation loss increases with range for spherical spreading (Eq. 6.5), cylindrical spreading (Eq. 6.6), and combined spherical/cylindrical spreading with a transition range of 100 m (Eq. 6.7). The effect of absorption (Eq. 6.8) in addition to spherical spreading is also shown for frequencies of 1, 10, and 100 kHz.

### 6.4.3 Additional Losses

#### 6.4.3.1 The Air–Water Interface

**Reflection and Transmission Coefficients**

In animal bioacoustics as well as noise research, one typically deals with sounds in one medium (i.e., either air or water) and then sticks to this medium, only modeling propagation within this medium and only considering receivers in this medium. However, sound does cross into other media, and so a fish might be able to hear an airplane flying overhead, and a bird flying directly overhead might be able to hear a submarine's sonar (Fig. 6.9).

As sound hits an interface, the incident wave, in most situations, gives rise to a reflected wave and a transmitted wave[4] (also see Chap. 5, where reflection is explained based on Huygens' principle). The energy of the reflected wave remains within the medium of the incident sound, but the energy of the transmitted wave is lost from the medium of the incident sound and transmitted into the adjacent medium. The amplitudes of the reflected and transmitted (plane) waves are given

[4] Dan Russell's animations of waves being reflected from hard and soft boundaries, and being transmitted: https://www.acs.psu.edu/drussell/Demos/reflect/reflect.html; accessed 12 October 2020.

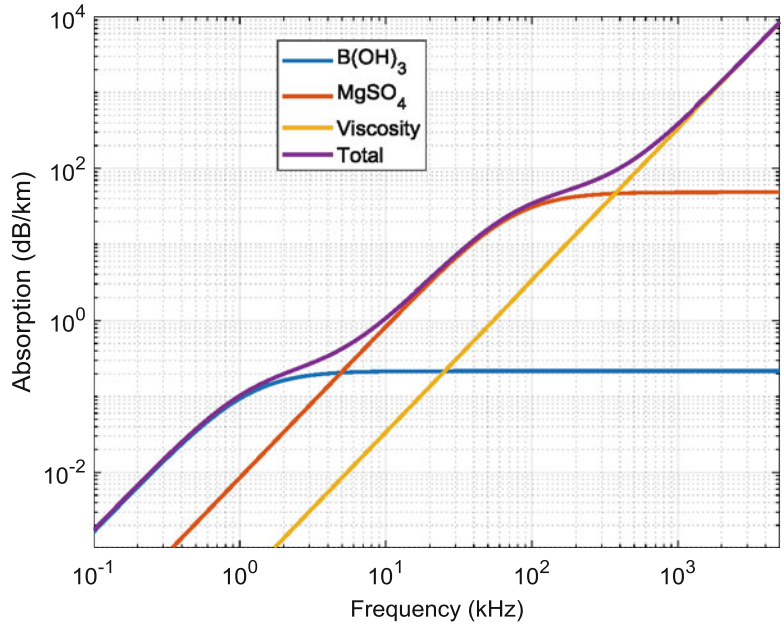



**Fig. 6.7** Graph of absorption loss dominated by $B(OH)_3$ for $f < 5$ kHz, by $MgSO_4$ for 5 kHz $< f <$ 500 kHz, and by viscosity above. $T = 10$ °C, $S = 35$ psu, $z = 0$ m, pH $= 8$

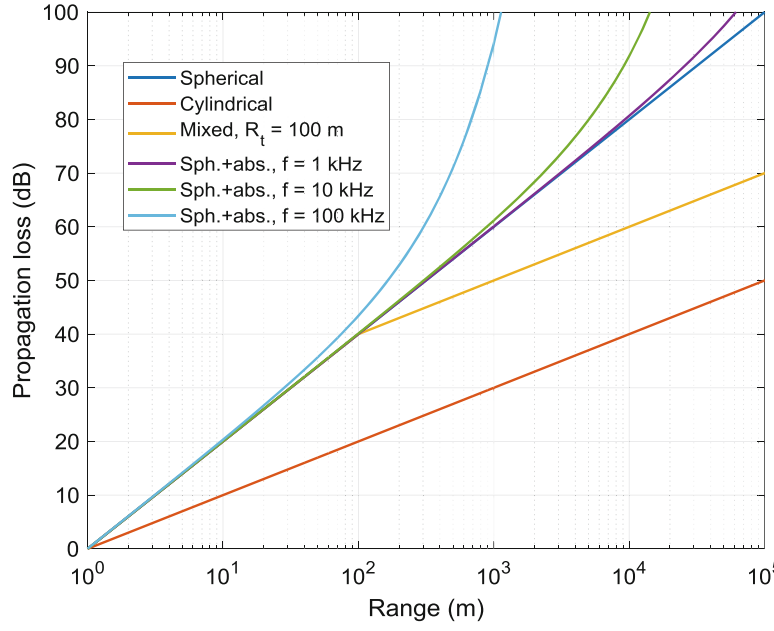



**Fig. 6.8** Plot of propagation loss versus range assuming spherical spreading (Eq. 6.5), cylindrical spreading (Eq. 6.6), and mixed spherical/cylindrical spreading (Eq. 6.7) for a transition range of 100 m. Propagation loss is also shown for spherical spreading with the addition of absorption (Eq. 6.8) corresponding to frequencies of 1, 10, and 100 kHz. Note that in the literature, the $y$-axis is sometimes flipped

by the reflection and transmission coefficients $\mathcal{R}$ and $\mathcal{T}$ (Medwin and Clay 1998):

$$\mathcal{R} = \frac{Z_2 \sin\theta_1 - Z_1 \sin\theta_2}{Z_2 \sin\theta_1 + Z_1 \sin\theta_2} \tag{6.10}$$

$$\mathcal{T} = \frac{2Z_2 \sin\theta_1}{Z_2 \sin\theta_1 + Z_1 \sin\theta_2}$$

where $\theta_1$ is the grazing angle of the incident wave, measured from the interface, and $\theta_2$ is the grazing angle of the transmitted (refracted) wave, also measured from the interface. The angle of incidence is measured from the normal (i.e., perpendicular to the interface); the angle of incidence and the grazing angle of the incident wave always add to 90°. The acoustic impedance $Z$ is the

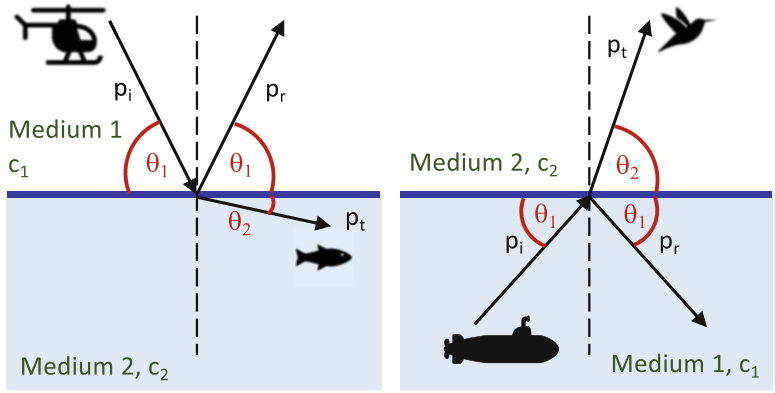



**Fig. 6.9** Sketches of a sound source in the air (helicopter; left) and water (submarine; right), and the incident $p_i$, reflected $p_r$, and transmitted $p_t$ rays (i.e., vectors pointing in the direction of travel, perpendicular to the wavefront), with corresponding grazing angles $\theta_1$ and $\theta_2$. In the left panel, medium 1 corresponds to air with sound speed $c_1$, and medium 2 corresponds to water with sound speed $c_2$. The situation is reversed in the right panel, where medium 1 is water, and medium 2 is air

product of density and sound speed: $Z = \rho c$. In air at 0 °C, $Z = 1.3$ kg/m$^3$ × 330 m/s = 429 kg/(m$^2$s). In freshwater at 5 °C, $Z = 1000$ kg/m$^3$ × 1427 m/s = 1,427,000 kg/(m$^2$s). In sea water at 20 °C and 1 m depth with 34 psu salinity, $Z$ = 1035 kg/m$^3$ × 1520 m/s = 1,573,200 kg/(m$^2$s) (see Chap. 4). So, $Z_{air} << Z_{water}$, whether it is freshwater or saltwater.

Snell's law (Fig. 6.9, Eq. 6.11)[5] relates the angles of the incident and refracted waves ($\theta_1$ and $\theta_2$) at the interface. Rays bend towards the interface, if the speed of sound in medium 2 is greater than that in medium 1 ($c_2 > c_1$) and away from the interface, if $c_1 > c_2$. While Snell's law typically relates the sines of the angles measured from the normal, it may also be expressed in terms of the cosines of the grazing angles (Etter 2018):

$$\frac{\cos\theta_1}{\cos\theta_2} = \frac{c_1}{c_2} \tag{6.11}$$

For normal incidence, all of the angles in Eq. (6.10) are 90°, and so all of the sines are 1, hence

$$\mathcal{R} = \frac{Z_2 - Z_1}{Z_2 + Z_1} \text{ and } \mathcal{T} = \frac{2Z_2}{Z_2 + Z_1}$$

For a sound source in air, $Z_1 << Z_2 => \mathcal{R} \rightarrow 1$ and $\mathcal{T} \rightarrow 2$, at normal incidence. Almost all of the sound is reflected, but the pressure in the water increases by a factor 2. The air–water boundary, for sound arriving from air, is considered "hard." The value of $\mathcal{T}$ is the reason why even weak aerial sources (such as drones hovering over whales) can be detected in water, below the source, at several meters depth (Erbe et al. 2017b), and commercial airplanes can be recorded in coastal waters, lakes, and rivers even if flying at hundreds of meters in altitude (Erbe et al. 2018). Received levels under water from airplanes may exceed behavioral response thresholds for underwater sound sources (Kuehne et al. 2020). For non-normal incidence, with $c_2 > c_1$, there exists a critical angle, beyond which the transmitted wave disappears. This situation is called total internal reflection. The only sound in the water is an evanescent field that decays exponentially in amplitude below the sea surface. The evanescent field is only important if the depth of the receiver is smaller than the in-water acoustic wavelength.

For a sound wave meeting the water–air interface from below, $Z_1 >> Z_2$ therefore $\mathcal{R} \rightarrow -1$ and $\mathcal{T} \rightarrow 0$. Almost all sound is reflected, albeit at

[5] Dan Russell's animation of refraction and Snell's law: https://www.acs.psu.edu/drussell/Demos/refract/refract.html; accessed 12 October 2020.

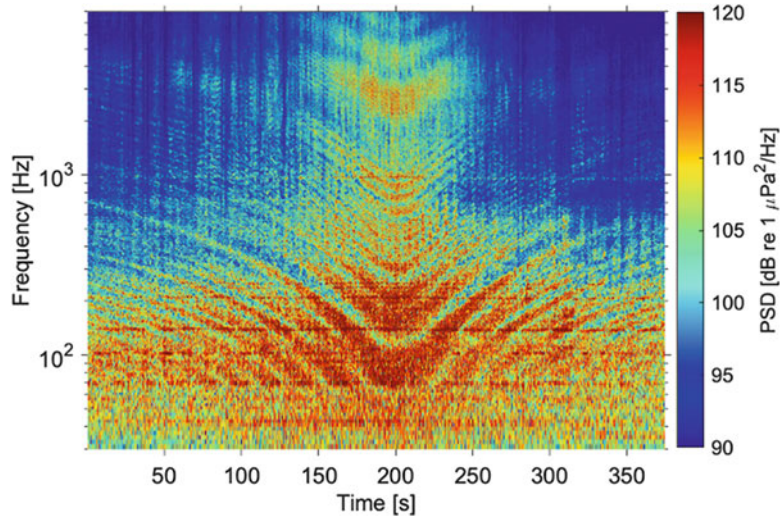



**Fig. 6.10** Spectrogram of the recording of a ship passing by a moored recorder, showing the pattern of constructive and destructive interference called the Lloyd's mirror effect. The closest point of approach occurred at about 200 s. Modified from (Erbe et al. 2016c); © Erbe et al. 2016; https://www.sciencedirect.com/science/article/pii/S0025326X15302125. Published under CC BY 4.0; https://creativecommons.org/licenses/by/4.0/

negative amplitude, which means that the incident and reflected pressures cancel each other out. This is why the water–air interface is called a pressure-release boundary (or "soft" boundary) for sound incident from below. For non-normal incidence, $\mathcal{R}$ and $\mathcal{T}$ need to be computed with Eq. (6.10). Also, as a sound source is moved to shallower depth (i.e., closer to the sea surface), the proportion of transmitted sound increases. This is because of the evanescent (i.e., exponentially decaying) field, which is ignored by Eq. (6.10), but that might still have enough amplitude at the sea surface for shallow sources (Godin 2008).

### Lloyd's Mirror

While not resulting in a loss of sound energy, the Lloyd's mirror effect is a result of reflection from the water–air interface from shallow sound sources. An omnidirectional source (i.e., one that emits sound in all directions) close to the sea surface (such as a ship's propeller) emits some of its sound in an upwards direction, and this sound reflects off the sea surface. At any receiver location, sound that traveled along the surface-reflected path overlaps with sound that traveled along the direct path from the source to the receiver. The reflected ray's amplitude is opposite in sign to the incident ray's amplitude ($\mathcal{R} = -1$); conceptually, this ray emerged from an image source (also called virtual source) with negative amplitude on the other side of the interface. The direct ray does not experience a flip in amplitude. Depending on the relative path lengths, the surface-reflected sound will add constructively to the sound that traveled along the direct path, or they will cancel each other out. This creates a pattern of constructive and destructive interference about the sound source, called the Lloyd's mirror effect. As a ship passes a moored recorder, the spectrogram shows the characteristic U-shaped interference pattern as successive peaks and troughs in amplitude at any one frequency over time (Fig. 6.10). Additional images of the Lloyd's mirror interference pattern can be found in (Parsons et al. 2020) for small electric ferries and in (Erbe et al. 2016b) for recreational swimmers and boogie boarders.

### Scattering at the Sea Surface

If the sea surface is not flat, then some of the reflected energy is scattered away from the geometric reflection direction, reducing the amplitude of the geometrically reflected wave. This is called surface scattering loss, which increases as the roughness of the sea surface increases, the acoustic wavelength decreases (i.e., acoustic frequency increases), and the grazing angle between the direction of the incident wave and the plane of the sea surface increases. This relationship is quantified by the Rayleigh roughness parameter (Jensen et al. 2011):

$$\gamma = 4\pi \frac{h}{\lambda} \sin\theta \tag{6.12}$$

where $h$ is the root-mean-square (rms) roughness of the surface (i.e., approximately ¼ of the significant wave height), $\lambda$ is the acoustic wavelength, and $\theta$ is the grazing angle. The larger the value of $\gamma$ is, the larger is the apparent roughness of the surface. The corresponding effective pressure reflection coefficient of the sea surface is then given by:

$$\mathcal{R}' = -e^{-0.5\gamma^2} \tag{6.13}$$

which corresponds to an additional propagation loss of $20 \log_{10}|\mathcal{R}'| = 4.34\gamma^2$ dB each time the sound reflects off the surface (Fig. 6.11). Note, however, that these formulae are only valid for surfaces that are not too rough, which, in this case, means $\gamma < 2$, corresponding to a scattering loss < 17 dB per bounce.

Strictly speaking, the effective pressure reflection coefficient (Eq. 6.13, Fig. 6.11) applies to the coherent component of the acoustic field, which can be thought of as the component that does not change as the rough sea surface moves. There will also be a scattered component that does change, and in some situations, this is an important contributor to the received signal. This component is ignored by Eq. (6.13), which can therefore be considered to provide an upper limit on the propagation loss per bounce.

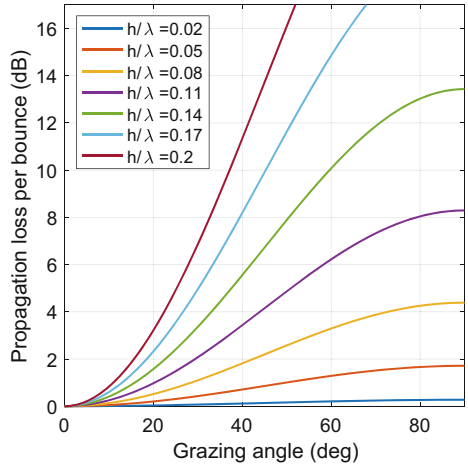



**Fig. 6.11** Graphs of additional propagation loss per bounce as a function of grazing angle for reflection from rough surfaces with various ratios of rms roughness to acoustic wavelength

### 6.4.3.2 The Seafloor Interface

The interaction of sound with the seafloor is more complicated. The acoustic properties of the seabed are often similar to those of the water, so a significant amount of sound can penetrate the seabed. The lower the frequency is, the deeper the sound can penetrate. At frequencies below a few kHz, it is common for a significant amount of acoustic energy to be reflected back into the water column from geological layering within the seabed. Seismic survey companies searching for oil and gas reserves are taking advantage of this.

Some of this complexity is illustrated in Fig. 6.12, which plots the pressure reflection coefficient as a function of grazing angle for four different seabed types: silt, sand, limestone, and basalt. Silt and sand layers are unconsolidated, which means that shear waves have a low speed and attenuate rapidly. (Shear waves are waves in which the particles oscillate at right angles to the direction of sound propagation; see Chap. 4.) Acoustically, they can often be well approximated by a fluid (which does not support shear waves at all) with an increased attenuation to account for the shear wave losses.

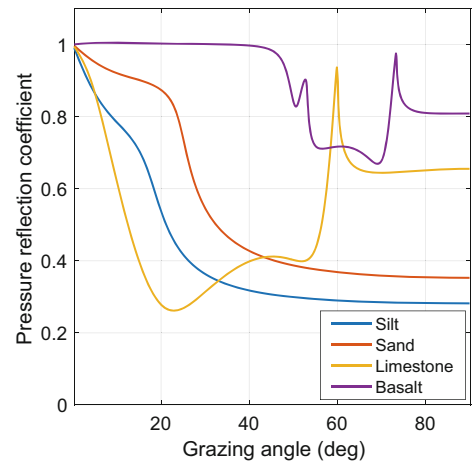



**Fig. 6.12** Curves of pressure reflection coefficient versus grazing angle for four different seabed types, calculated with parameters from Jensen et al. (2011)

Unconsolidated sediments become more reflective as the sediment grain size increases from silt to sand. Limestone and basalt are consolidated rocks, which allow both compressional waves and shear waves to propagate, and are thus referred to as solid elastic seabeds. Basalt is a hard rock and highly reflective at all grazing angles. The reflection coefficient of limestone, however, is perhaps surprising. While it is also a rock, it has the lowest reflectivity of the four seabeds at small grazing angles. This is because the shear wave speed in limestone is very similar to the sound speed in water, which allows energy to pass easily from sound waves in the water to shear waves in the seabed.

Curves of reflection coefficients versus grazing angle are even more complicated for layered seabeds due to interference between waves reflecting from different layers, and in this case, the reflectivity becomes frequency dependent. Despite the complexity, there are computer programs available, based on techniques described in Jensen et al. (2011), that can numerically calculate the reflection coefficient curve for any arbitrarily layered seabed. A good example is *BOUNCE*, which is part of the *Acoustics Toolbox*.[6] A much bigger problem is the common lack of information on the geoacoustic properties of the seabed, to be able to provide these programs with accurate input data.

Seafloor roughness can further reduce the apparent acoustic reflectivity, although if the rms roughness is known, this can be dealt with (at least approximately) by using Eq. (6.12) to calculate the associated Rayleigh roughness parameter $\gamma$ as a function of grazing angle. The effective seabed reflection coefficient is then:

$$\mathcal{R}' = \mathcal{R} e^{-0.5\gamma^2} \tag{6.14}$$

where $\mathcal{R}$ is the pressure reflection coefficient for the flat seafloor (Eq. 6.10). All terms in this equation depend on grazing angle. The propagation loss per bounce is given by $20 \log_{10} |\mathcal{R}'|$.

#### 6.4.3.3 Scattering Within the Water Column

Sound can be scattered within the water column by anything that causes sharp changes in sound speed, density, or both (i.e., acoustic impedance, which is the product of sound speed and density; see Chap. 4). This includes gas bubbles, biological organisms (in particular those with gas-filled organs like lungs or swim bladders), and suspended sediment particles. Water column scattering is utilized in active sonar systems, which rely on the backscattered signal to detect and/or characterize objects within the water column. However, clouds of air bubbles formed by breaking waves can cause an appreciable increase in propagation loss in some circumstances.

Air bubbles are essentially small, resonant cavities within the water column, which can both scatter and absorb sound and, when found in large numbers, can change the effective density, and hence sound speed, of the water. When a wave breaks, it entrains a large amount of air down to depths of several meters, forming a cloud of bubbles of a range of sizes. The large bubbles rise to the surface quite quickly, but the smaller bubbles can remain at depth for many minutes. This can increase the propagation loss for sound traveling close to the surface (Ainslie 2005; Hall 1989).

### 6.4.4 Numerical Propagation Models

#### 6.4.4.1 The Wave Equation and Solution Approaches

The ocean is a complicated environment for sound propagation, and the simple approaches to estimating propagation loss described above are very limited in their applicability. As a result, a great deal of effort has gone into developing numerical propagation models that can calculate acoustic propagation loss for realistic situations. What follows is a brief introduction to the topic. The interested reader is referred to Etter (2018)

[6] *Acoustics Toolbox:* https://oalib-acoustics.org/models-and-software/acoustics-toolbox/; accessed 30 September 2020.

and Jensen et al. (2011) for a more comprehensive treatise.

Fundamentally, all numerical propagation models solve the acoustic wave equation, which is a differential equation that relates the way the pressure changes over time to how it changes spatially as a wave propagates:

$$\nabla^2 \Phi = \frac{1}{c^2} \frac{\partial^2 \Phi}{\partial t^2} \tag{6.15}$$

where $\nabla^2$ is the Laplace operator, $\partial$ indicates the partial derivative, $c$ is the speed of sound, $t$ represents time, and $\Phi$ is the solution to the wave equation.

The wave equation itself is well understood and straightforward to solve in simple cases; however, there are two issues that make it difficult to solve numerically for typical underwater acoustics problems:

1. Solutions are usually desired over domains that are orders of magnitude larger than the acoustic wavelength. Direct solution methods, such as finite differences or finite elements, require meshing the solution domain at a resolution of a small fraction of a wavelength, so the size of the required domain makes these approaches impractical for most propagation problems, even with modern computing hardware.
2. The boundaries of the domain, particularly the seabed, are complicated, but very important to model accurately as they have a strong influence on sound propagation.

Getting around these difficulties requires making approximations that lead to equations that are practical to solve for the problems of interest, with different approximations leading to different methods suitable for different situations.

In general, the solution of the acoustic wave equation is a function of three spatial dimensions and time. In Cartesian coordinates, the acoustic pressure can be written as: $p(x, y, z, t)$. In most cases, we are interested in the field generated by a small source, which can be approximated as a single point in space. It is more convenient to work in cylindrical coordinates centered on the source location, $p(r, z, \phi, t)$, where $r$ is the horizontal distance from the source to the receiver, $z$ is the receiver depth below the sea surface, and $\phi$ is the horizontal plane azimuth angle of the receiver relative to some direction reference.

Many modeling approaches start by assuming that the solution has a harmonic time dependence so that $p(r, z, \phi, t) = p_\omega(r, z, \phi)e^{-i\omega t}$ where $\omega = 2\pi f$ is the angular frequency and $i = \sqrt{-1}$. Substituting this solution form into the wave equation (Eq. 6.15) leads to another differential equation called the Helmholtz equation, which can be solved at a specified $\omega$ to give $p_\omega(r, z, \phi)$. The computational advantage of this is that the Helmholtz equation can be solved independently for each required frequency, converting a coupled four-dimensional (4D) problem into a number of independent 3D problems. Models that use this approach are known as frequency domain models, whereas models that directly solve the wave equation are known as time domain models. If required, the time domain solution can be reconstructed from multiple frequency domain solutions using Fourier synthesis (see Jensen et al. 2011, Chap. 8, for details).

The azimuth angle dependence can be dealt with by two different approaches. Modeling in 3D retains the full azimuth dependence of the environment, whereas $N \times$ 2D modeling assumes that changes in the environment due to small changes in $\phi$ have negligible effect on sound propagation, so that modeling can be carried out independently along each azimuth of interest. The majority of numerical models use the $N \times$ 2D approach, because there is again a substantial computational saving, this time by reducing a coupled 3D problem, solving for $p_\omega(r, z, \phi)$, to a number of independent 2D problems, each solving for $p_{\omega, \phi}(r, z)$ using only environmental information for the corresponding azimuth.

The inherent assumption of the $N \times$ 2D method provides a good approximation to the sound field in many propagation modeling situations where horizontal sound speed gradients are much smaller than vertical sound speed gradients, the seabed slopes are small, and the ranges are not large enough for the remaining out-of-plane effects to have an appreciable effect

on the sound field. However, there are cases where full 3D modeling may be required; for example, around steep-sided submarine canyons, in the presence of nonlinear internal waves that can produce strong horizontal sound speed gradients, or for very-long-range propagation across ocean basins.

Some propagation models further simplify their calculations by assuming that the environment (but not the sound field) is independent of range, which means that the sound speed profile is a function of depth only, and the water depth and seabed properties are the same at all ranges (i.e., the seafloor is flat). These are called range-independent (RI) propagation models, whereas propagation models that allow the sound speed profile and/or the water depth and/or the seabed properties to vary with range are known as range-dependent (RD) models.

Acoustic propagation models are usually characterized by the numerical approach adopted, and the following sections described some of the most common. Guidance on which propagation model to use in various scenarios follows this section.

#### 6.4.4.2 Ray and Beam Tracing

A ray is a vector, normal to the wavefront, and shows the direction of sound propagation. Ray models trace rays by repeatedly applying Snell's law (Eq. 6.11). For layered media (such as layers of ocean water with differing properties), Snell's law relates the angles of incidence $\theta_1$ and refraction $\theta_2$ at every layer boundary. Rays bend towards the horizontal, if $c_2 > c_1$, and away from the horizontal if $c_1 > c_2$.

There are several approaches to calculating the amplitude of the acoustic field. The simplest, known as conventional ray tracing, is to use the distance between initially adjacent rays to determine the area over which the sound power has spread and calculate the intensity as the power per unit area. Unfortunately, this method results in unphysical predictions of infinite sound amplitude at locations called caustics, where initially adjacent rays cross and therefore have zero separation. It also predicts sharp transitions to zero sound intensity in shadow zones, which are regions where rays do not enter, whereas in reality, the transition will be smoother. Both of these problems are a result of a high-frequency approximation inherent in ray theory, which cannot deal with diffraction (i.e., the phenomenon of waves bending around obstacles or spreading out after passing through a narrow gap; see Chap. 5 on sound propagation examples in the terrestrial world).

An alternative approach to calculating the amplitude of the acoustic field is to treat each ray as the center of a beam with a specified (usually Gaussian) amplitude profile. The field at a particular location is then obtained by summing the contributions from all the beams that overlap at that location. The main challenge with this approach is determining how the amplitude and width of the beam should change along the ray, but algorithms have been developed to do this (see Jensen et al. 2011, Sect. 3.5, for details). One of the best-known propagation codes of this type is *Bellhop* (Porter and Bucker 1987), a fully range-dependent, Gaussian beam tracing program suitable for $N \times 2$D modeling that is available as part of the *Acoustics Toolbox*. The toolbox also includes a fully 3D variant called *Bellhop3D*.

Although Gaussian beam tracing is an improvement to conventional ray tracing and reduces the effects of the high-frequency assumption inherent in ray theory, it does not completely eliminate them. Its treatment of shadow zones and caustics produces realistic, but not necessarily accurate results and, importantly, it does not predict waveguide cutoff effects.

In underwater acoustics, the term waveguide or duct is used to describe any situation in which sound is constrained to a particular span of depths by reflection, refraction, or some combination of the two. Common examples include (Fig. 6.13):

1. A shallow-water duct in which sound is constrained by reflection from both the sea surface and the seabed.
2. A surface duct, in which the sound speed near the sea surface increases with increasing depth. This results in sound that is initially heading downward being refracted upwards towards the sea surface, where it is reflected back downward again, and so on. It is therefore constrained by reflection at the top and by

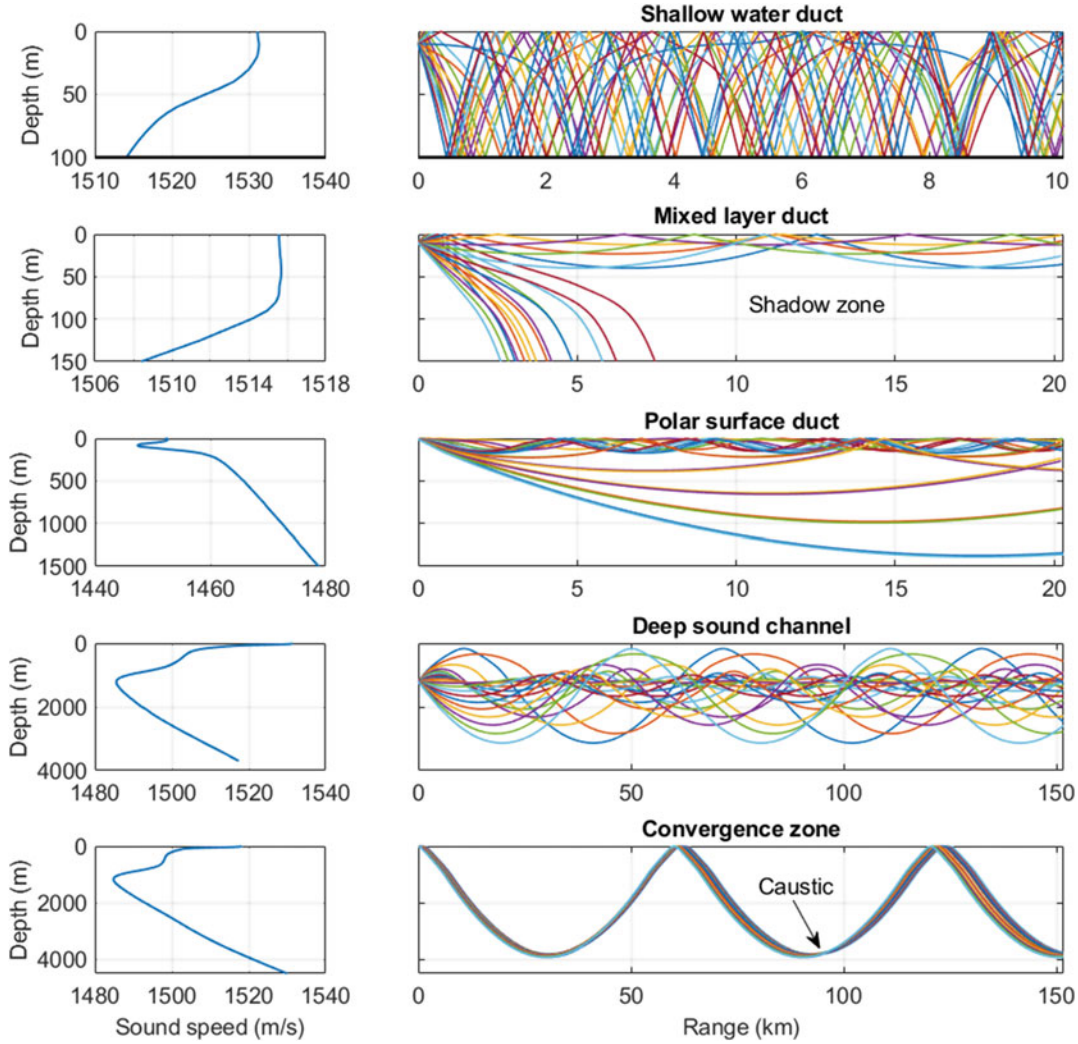



**Fig. 6.13** Sound speed profiles (left) and ray trace plots computed using *Bellhop* (Porter and Bucker 1987, right) illustrating the common underwater acoustic ducts described in the text. The source depth was 10 m for all except the deep sound channel example, which had a source depth of 1200 m

refraction at the bottom. Weak surface ducts are often found in the mixed layer due to sound speed increasing with increasing pressure, and strong surface ducts are ubiquitous in polar oceans because both pressure and temperature increase with increasing depth. Sea ice can, however, reduce the acoustic reflectivity of the sea surface and therefore increase the attenuation of sound traveling in the duct.

3. The Deep Sound Channel (DSC), also known as the sound fixing and ranging (SOFAR) channel, in which sound is refracted towards the minimum in the sound speed (i.e., towards the waveguide axis). The waveguide axis occurs at a depth of about 1000 m in much of the world's ocean. The sound is constrained by refraction both above and below the axis of the waveguide. However, these are not sharp boundaries, and the steeper the angle of propagation is, the larger are the excursions of the ray paths away from the axis.
4. Convergence zone propagation in which sound is constrained by reflection from the sea surface and refraction from the increase

of sound speed with increasing depth that occurs below the axis of the DSC.

In all cases, the waveguide will only trap rays leaving the source within a certain span of angles from the horizontal. In the case of the shallow water waveguide, this is because the seabed reflectivity reduces as the grazing angle increases (Fig. 6.12), so more energy is lost on each bottom bounce at steeper angles. In the other waveguide cases, it is because the refraction is not strong enough to turn the ray around before it either reaches a depth where the sound speed gradient is refracting it away from the waveguide (surface duct) or it hits the seabed (DSC and convergence zone).

According to ray theory, rays can be launched at any angle, irrespective of the frequency, and so it should always be possible to find rays that will be trapped in the waveguide, provided the source is at a suitable depth. However, this is not actually the case at low frequencies, where the acoustic wavelength becomes an appreciable fraction of the thickness of the waveguide. It turns out that if the frequency is sufficiently low, no energy will be trapped in the waveguide, and the waveguide is said to be cut off. Understanding why this is the case requires an understanding of normal modes, which is the topic of the next section.

#### 6.4.4.3 Normal Modes

Most people find the concept of normal modes to be less intuitive than that of rays, but it is very useful for understanding low-frequency sound propagation in the ocean and forms the basis for a class of acoustic propagation models called normal-mode models.

Normal modes are best understood by first considering an ideal shallow-water waveguide with a constant depth (i.e., flat seafloor), constant sound speed, and perfectly reflecting seafloor. Solving the Helmholtz equation for this situation requires that two so-called boundary conditions be met: one at the sea surface and one at the seafloor. The sea surface is a soft boundary as far as underwater sound is concerned, so the boundary condition here is that the acoustic pressure due to the incident and reflected waves sums to zero, which requires that an incident sound wave is inverted on reflection. Conversely, the seafloor is a hard boundary, which requires that the incident and reflected waves sum to a maximum pressure; so the amplitudes of the incident and reflected waves must have the same sign.

Both of these boundary conditions have to be satisfied simultaneously. The water depth is fixed, and normal modes consider one frequency at a time, so the wavelength is fixed. The only variable that can change to satisfy the requirements is the angle from the horizontal at which the wave propagates. There are certain, discrete propagation angles that allow the surface and seafloor boundary conditions to be met simultaneously, corresponding to the normal modes. Each normal mode consists of a pair of plane waves, one propagating upward and the other downward, at the same angle to the horizontal (Fig. 6.14). The mode that corresponds to the pair of waves propagating closest to the horizontal is called the lowest-order mode (mode 1), and the mode order increases as the propagation angle gets steeper. Note that the waves can never propagate exactly horizontally, because that does not meet the boundary conditions.

A receiver in the water column will receive the sum of the pressures from the upward and downward traveling waves. The amplitude of that combined signal can be plotted as a function of depth and range for each mode, yielding a series of mode shape curves (Fig. 6.15). Note that there is always a null in pressure (i.e., a node) at the sea surface and a maximum in pressure magnitude (i.e., $+1$ or $-1$; an antinode) at the hard seafloor.

The mode shapes are reminiscent of standing waves on a guitar string, which are also normal modes. However, on a guitar string, different modes correspond to different frequencies of vibration, whereas in a waveguide, different modes correspond to sound of the same frequency propagating at different angles to the horizontal.

For any waveguide thickness, the propagation angles for a particular mode increase as frequency is reduced. The ideal waveguide considered so far has no limit to how steep the propagation angles can be, but that is not the case for real ocean

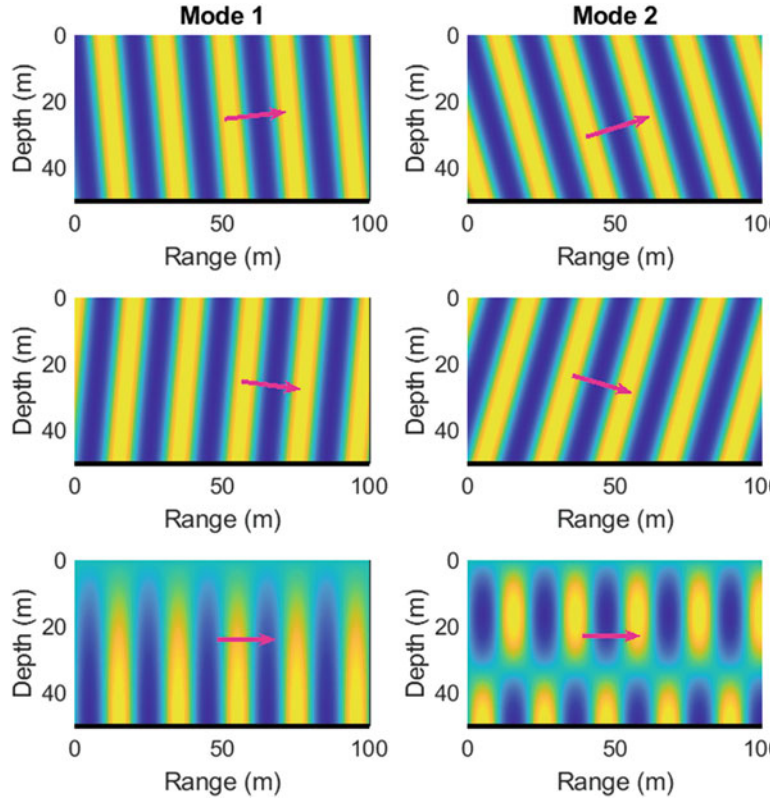



**Fig. 6.14** Depth-range plots showing how the normal modes of an ideal shallow-water waveguide (lower panel) result from a pair of upward (upper panel) and downward (middle panel) propagating plane waves. Left-hand panels are for mode 1, right-hand panels are for mode 2. Arrows show the direction of propagation. The water depth is 50 m and the acoustic wavelength is 20 m

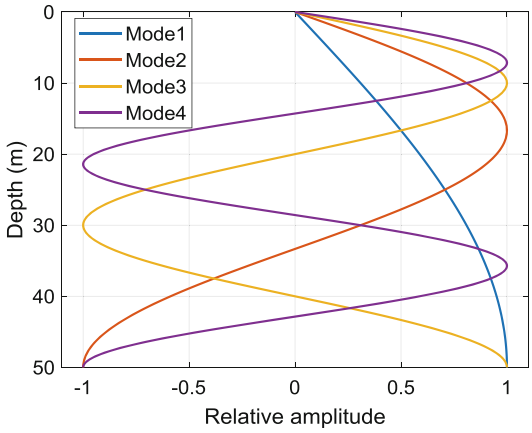



**Fig. 6.15** Mode shapes for the first four normal modes of a 50-m deep ideal shallow-water waveguide with a rigid seabed

waveguides which, as discussed in the previous section, all have limits on the angular range of the energy they can trap. The highest-order mode corresponds to the steepest propagation angle, so as frequency is reduced, it will become too steep to be constrained by the waveguide and will no longer be able to propagate. As frequency is reduced further, the same will happen to the next-highest-order mode, and so on until the lowest-order mode is unable to propagate, at which point the waveguide is said to be cut off.

In real ocean waveguides, the sound speed varies with depth, which causes the propagation angle of each mode to also be a function of depth. This changes the mode shapes, but you can still consider a mode to consist of a pair of upward and downward going waves, propagating at the same angle to the horizontal at any given depth.

The starting point for the mathematical derivation of normal-mode models is the depth-separated Helmholtz equation, which is valid for range-independent problems and is obtained by assuming that the acoustic field can be represented by the product of a function of depth and a function of range:

$$p_{\omega,\phi}(r,z) = F(z)G(r).$$

Substituting this into the Helmholtz equation results in a one-dimensional differential equation for $F(z)$ in terms of a separation constant $k_r$. The solution of this differential equation has poles (infinities) at certain values of $k_r$, which correspond to the normal modes. Normal-mode codes search for these values of $k_r$, calculate the corresponding mode shapes, and then compute $p_{\omega,\phi}(r, z)$ by a mathematical technique called the "method of residues," which involves summing the contributions of all the poles, which in this case, corresponds to summing the contributions of the individual modes. It turns out that $k_r$ has a geometric interpretation. It is called the horizontal wavenumber and is related to the modal propagation angle $\theta$ (relative to the horizontal) by $k_r = \omega \cos(\theta)/c$.

Normal-mode codes are computationally very fast for range-independent problems, because the modes only have to be found once, after which the field can be calculated at any desired range with very little additional computational effort.

Dealing with range-dependent problems involves approximating the environment as a series of range-independent sections, calculating the modes for each of these sections, and then calculating how the energy present in the modes in one section transmits across the boundary to the modes in the next section. There are two approaches:

1. The adiabatic mode method assumes that all the energy in mode 1 stays in mode 1, all the energy in mode 2 stays in mode 2, etc. This is relatively simple to implement and fast to compute, but is only accurate for environments that change relatively slowly with range.
2. The coupled-mode method allows energy to transition between modes, and so can deal with environments that change more rapidly. But this method is much more computationally demanding.

A good example of a normal-mode model is *KRAKEN* (Porter and Reiss 1984), which can be used for both range-independent and range-dependent modeling (both adiabatic and coupled) and is part of the *Acoustics Toolbox* (Footnote 5).

One limitation of normal-mode models such as *KRAKEN* is that they only include the component of the acoustic field that is fully trapped in the waveguide, so they tend to be inaccurate at short ranges where the component of the field that is losing energy out of the waveguide can be significant. This problem can be addressed by including so-called leaky modes in the solution. However, reliably finding leaky modes turns out to be a very challenging numerical task. The most successful normal-mode model to-date in this respect is *ORCA* (Westwood et al. 1996), which is accurate at short range and can also deal with seabeds that support shear waves. *ORCA* was written as a range-independent model, but there have been several attempts to adapt it to range-dependent problems using the adiabatic mode method (Hall 2004; Koessler 2016).

#### 6.4.4.4 Wavenumber Integration

The mathematical derivation of the wavenumber integration method also starts with the depth-separated Helmholtz equation, but in this case, $F(z)$ is calculated by direct numerical solution of the one-dimensional differential equation over a range of $k_r$ values, giving the so-called wavenumber spectrum. The acoustic field $p_{\omega,\phi}(r, z)$ is then obtained by an integral transform of the wavenumber spectrum that involves a Hankel function. A numerical approximation to the Hankel function that is valid except at ranges smaller than the acoustic wavelength can be used to convert this integral transform into a Fourier transform, which can then be evaluated using the very efficient Fast Fourier Transform algorithm.

Wavenumber integration codes that use this method of evaluating the integral transform are known as fast-field programs. Common examples are *SAFARI*, *OASES*, and *SCOOTER* (Porter 1990; Schmidt and Glattetre 1985). *OASES* is a development of *SAFARI* and has largely superseded it, whereas *SCOOTER*, which is part of the *Acoustics Toolbox* (Footnote 5), is a separate, but largely equivalent, development. These programs are very accurate for acoustic propagation calculations at ranges close enough to the source that the environment can be considered range-independent, and can deal with arbitrarily complicated, layered seabeds. For

most applications, the short-range limitation introduced by the Hankel function approximation is of little consequence, but, if necessary, it can be removed (at additional computational cost) by directly evaluating the integral transform.

It has proved difficult to extend the wavenumber integration method to range-dependent problems in a way that results in an efficient propagation model, although the full (paid) version of *OASES*[7] does have this capability. The theoretical background of this model is described in Goh and Schmidt (1996).

#### 6.4.4.5 Parabolic Equation

Inserting a solution of the form $p_{\omega,\phi}(r,z) = f(r,z)H_0^{(1)}(k_0 r)$ into the Helmholtz equation yields parabolic-equation (PE) models. Here, $H_0^{(1)}$ represents an outgoing cylindrical wave with wavenumber $k_0 = 2\pi f/c_0$ where $c_0$ is an assumed sound speed. Technically, $H_0^{(1)}$ is a Hankel function of the first kind of zero order. The aim of PE models is to solve for $f(r, z)$, which represents the way in which the true field varies from that produced by the ideal outgoing cylindrical wave.

If the sound is assumed to be propagating predominantly in the range direction (the so-called paraxial approximation), then an efficient numerical algorithm can be employed. Given $f(r, z)$, a small range step $dr$ is added to calculate $f(r + dr, z)$, a little bit farther from the source. This calculation can then be repeated as many times as desired to march the solution out in range. The sound field at one range is thus used to calculate the sound field at the next range and so on, without explicitly solving the depth-separated Helmholtz equation, making this a fundamentally different approach to the normal mode and wavenumber integration methods discussed previously.

Initially, the paraxial approximation was very restrictive and severely limited the utility of PE models for solving underwater acoustics problems. The more recent development of so-called high-angle PE models greatly relaxed this approximation. The way in which the solution marches out in range makes it straightforward to include range-dependent water depth, sound speed profiles, and seabed properties, and as a result, high-angle PE models have become the method of choice for solving range-dependent propagation problems.

Perhaps the most widely used PE model is *RAM* (Collins 1993), which allows the user a trade-off between the valid angular range and computational efficiency by specifying the number of terms to be used in a Padé approximation, which is central to the wide-angle algorithm. The more terms that are used in the Padé approximation, the wider is the valid angular range. Even though this allows the paraxial approximation to be greatly relaxed, it cannot be completely eliminated, and so PE models should always be used with care when acoustic energy propagating at steep angles is significant.

Another consideration when running *RAM* or similar PE models is that they use a finite computational grid in the depth direction, and energy will be artificially reflected by the sudden truncation at the bottom of the grid. This is usually dealt with by including an extra attenuation layer underneath the layer representing the physical seabed. The attenuation layer has the same density and sound speed as the seabed but an artificially high attenuation coefficient so that little energy reaches the bottom of the grid, and any energy that does reflect is further attenuated before reappearing in the water column. A sudden change in attenuation can also lead to reflections, so in critical situations, it is advisable to ramp the attenuation up smoothly from its seabed value to a high value, rather than having a step change.

There are several variants of *RAM* intended for different purposes (Table 6.1). The only one that can deal with elastic seabeds is *RAMS*, but it requires careful tuning of parameters to avoid instability, and in some cases involving layered seabeds, it is impossible to obtain a stable solution. More recent PE models have been developed that overcome these limitations (Collis et al. 2008) yet are research codes not readily

[7] *OASES* code https://oceanai.mit.edu/lamss/pmwiki/pmwiki.php?n=Site.Oases; accessed 1 October 2020.

**Table 6.1** Summary of variants of the *RAM* parabolic-equation codes

| Program | Seabed layering | Seabed type | Sea surface |
|---|---|---|---|
| *RAM* | Specified relative to the sea surface. Bathymetry cuts through the stack of layers. | Fluid only | Flat |
| *RAMSurf* | As for RAM. | Fluid only | Specified profile |
| *RAMGeo* | Specified relative to the seabed. Layering follows bathymetry. | Fluid only | Flat |
| *RAMS* | As for RAM. | Elastic | Flat |

available. The majority of PE codes are intended for $N \times$ 2D modeling. However, research-level 3D PE codes have been developed (see Jensen et al. 2011, Sect. 6.8, for details).

### 6.4.5 Choosing the Most Appropriate Model

If the frequency is high enough that the acoustic wavelength is less than a small fraction of the smallest significant feature in the sound speed profile (e.g., mixed layer thickness, water depth), then use a ray tracing or beam model (e.g., *Bellhop*), otherwise use one of the low-frequency models. A rule of thumb for the 'small fraction' is 1/100. However, accurately modeling sound propagation in a weak duct may require the use of a low-frequency model up to a higher frequency than this rule would suggest. If in doubt, run some tests using both types of models to determine the frequency at which the two models start to agree.

When choosing a low-frequency model, if the range is short enough that the environment can be considered range-independent, then pick a wavenumber integration model (e.g., *OASES* or *SCOOTER*), otherwise use a PE model (e.g., *RAM*). The benefit of wavenumber integration for range-independent modeling is its greater accuracy at short range compared to either a normal-mode model (which only considers trapped energy) or a PE model (which has high-angle limitations). Wavenumber integration can also deal accurately with elastic seabed effects, which tend to be most important at short range. PE codes have largely replaced normal-mode codes for range-dependent modeling because of the greater practicality of the PE range-marching algorithm.

Range-dependent modeling with layered elastic seabeds remains a difficult computational task. One commonly resorts to work-around strategies, such as replacing the true seabed with an "equivalent" fluid seabed that has a similar reflection coefficient versus grazing angle dependence at low grazing angles. This allows a standard PE code to be used for the modeling but is only accurate at ranges large enough that there is no high-angle energy reaching the receiver.

### 6.4.6 Accessing Acoustic Propagation Models

Many of the models described in this chapter are freely available for download from the Ocean Acoustics Library[8] (OALIB). OALIB includes Michael D. Porter's *Acoustics Toolbox*, which incorporates a Gaussian beam tracing model (*Bellhop*), wavenumber integration code (*SCOOTER*), normal-mode model (*KRAKEN*), as well as several other useful programs including one for calculating seabed reflectivity as a function of grazing angle for arbitrarily complicated, layered seabeds (*BOUNCE*). These all use similar input and output file formats, have been regularly updated until at least 2020, and are well documented. A number of *MATLAB* (The MathWorks Inc., Natick, MA, USA) routines for dealing with the input and output are also provided. Also available on OALIB is the free version of the wavenumber integration code

[8] Ocean Acoustics Library https://oalib-acoustics.org/; accessed 17 June 2020.

*OASES* and a number of different PE codes, including the *RAM* family.

Unfortunately, downloading a particular code is often just the start of a journey that may include compiling it for the particular operating system you are using, deciphering the documentation to determine what input files are required and how they need to be formatted, and then working out how to read and plot the output data. There are usually a number of adjustable parameters that affect how the program operates, and it is necessary to have an understanding of the underlying numerical methods in order to set these appropriately. Inappropriate parameter selection will often lead to meaningless results, so whenever you start using a different propagation model, you should run a series of tests on simple problems (to which the answer is known) in order to make sure you are getting the correct results. The standard of documentation varies considerably between the different models that are available from OALIB and is minimal for some.

*AcTUP*[9] is a *MATLAB* GUI to earlier (2005) versions of the *Acoustics Toolbox* and several of the *RAM* family of PE codes. *AcTUP* comes packaged with the required Windows executables. This provides a convenient entry point for those new to acoustic propagation modeling as it allows different codes to be run on the same problem with minimal changes. However, careful parameter selection is still required in order to get meaningful results; put garbage in, get garbage out.

## 6.5 Practical Acoustic Modeling Examples

Having worked through the theory and concepts, this section finally puts all of the above into action and provides examples of some practical acoustic propagation modeling tasks of increasing complexity. These all involve the estimation of received levels due to a source with known sound emission characteristics, and are conceptually based on re-arranging the passive sonar equation (Eq. 6.1) to solve for the received level *RL*:

$$RL = SL - PL. \tag{6.16}$$

The tasks are:

1. Calculate *RL* as a function of range and depth in a given direction from a tonal (i.e., single-frequency) source.
2. Calculate *RL* as a function of range and depth in a given direction from a broadband source.
3. Calculate *RL* as a function of geographical position and depth for an omnidirectional source in a directional environment.
4. Calculate *RL* as a function of geographical position and depth for a directional source in a directional environment.

Indicative execution times are given for calculations that were carried out on a desktop computer with an Intel i7–7700 CPU, a clock speed of 3.6 GHz, and 64 GB of RAM. The processor had 4 physical cores but the models used here were single-threaded so only used one core. The computer was running a 64-bit Windows 10 operating system.

### 6.5.1 Received Level Versus Range and Depth from a Tonal Source

For this case, it is only necessary to specify the acoustic environment (i.e., bathymetry profile, sound speed profile, and seabed properties) along a single azimuth from the source. The propagation loss *PL* is only required at the source transmission frequency, and can be obtained using a single run of an appropriate propagation model. The received level *RL* can then be obtained using Eq. (6.16).

The example of a fin whale (*Balaenoptera physalus*) located about 40 km off the coast of southwestern Australia, at a depth of 50 m, while emitting a 20-Hz tone at a source level of 189 dB re 1 µPa m (Sirovic et al. 2007) is depicted in Fig. 6.16. The modeled direction of propagation

[9] *AcTUP* http://cmst.curtin.edu.au/products/underwater/download/; accessed 1 October 2020.

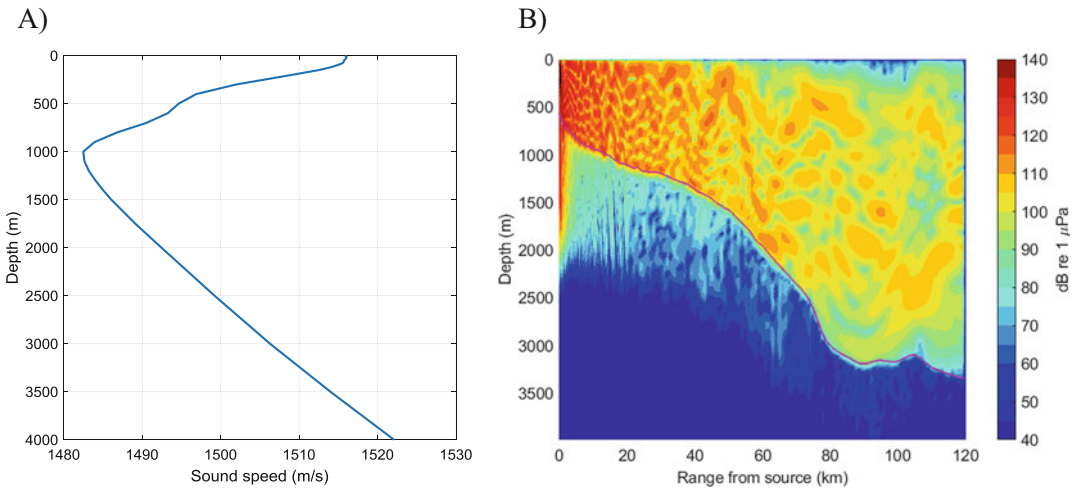



**Fig. 6.16** (**a**) Sound speed profile used for the modeling examples. (**b**) Modeled received SPL as a function of range and depth for a fin whale at a depth of 50 m emitting a 20-Hz tone with a source level of 189 dB re 1 μPa m. The magenta line is the seafloor

was due west from the source, and the bathymetry profile (i.e., magenta line in Fig. 6.16b) was interpolated from the Geosciences Australia 0.15′ resolution bathymetry database.[10] The sound speed profile (Fig. 6.16a) was calculated from salinity and temperature data obtained from the *World Ocean Atlas* (Locarnini et al. 2018; Zweng et al. 2018). The seabed was modeled as a fine sand half-space with parameters from Jensen et al. (2011). Propagation loss modeling was carried out with *RAMGeo* in *AcTUP*, which is very efficient at such a low frequency, taking only a few seconds. A simple program was written in *MATLAB* to read the propagation loss file produced by *RAMGeo*, calculate the received levels using Eq. (6.16), and plot the results. Note that *AcTUP* can be used to plot propagation loss, but not received level.

The sound field has a complicated structure of peaks and nulls that is the result of constructive and destructive interference between sound that has traveled from the source to the receiver via different paths. This is typical of the sound fields produced by tonal sources. The overall reduction in received level with increasing range is quite slow, particularly beyond 70 km, due to the sound becoming constrained by refraction in the deep sound channel. This is typical of downslope propagation from a near-surface source situated over the continental slope into deep water.

### 6.5.2 Received Level Versus Range and Depth from a Broadband Source

Many sources of underwater sound are broadband, which means that they produce significant acoustic output over a wide range of frequencies. Ships, pile driving, and the airgun arrays used for seismic surveying all produce broadband noise, and modeling the resulting sound fields is of importance when assessing the potential impacts of these sources on marine animals.

A common way to carry out broadband modeling for continuous sound such as ship noise is:

[10] Whiteway, T., Australian Bathymetry and Topography Grid, June 2009, https://ecat.ga.gov.au/geonetwork/srv/eng/catalog.search#/metadata/67703; accessed 6 November 2020.

1. Break the required frequency span into a series of frequency bands (e.g., 1/3 octave bands are commonly used; see Chap. 4).
2. Use a propagation model to estimate a typical propagation loss for each band. This can either be done by running the propagation loss model at the center frequency of each band or by running it at a number of frequencies within the band and then averaging the results. The latter is preferred as it smooths out the interference field to some extent, but if the source emits a wide range of frequencies that span many bands, then the two methods will yield very similar results for the total field.
3. Integrate the source power spectral density over each band and convert to a source level.
4. Use Eq. (6.16) to obtain the received level in each band.
5. Sum the corresponding mean-square pressures across the bands to obtain an overall mean-square pressure that can then be converted to an overall received sound pressure level (SPL, see Chap. 4).

The use of mean-square pressure as a metric is problematic for impulsive sources such as airguns or pile driving, because the results become very sensitive to the duration of the signal, which is often hard to determine. Source and received levels for impulsive sources are therefore usually characterized in terms of sound exposure, and its logarithmic measure, the sound exposure level (SEL, see Chap. 4).

Computing the received levels for impulsive sources follows the same steps as for broadband, continuous sources, except that in step 3, the source spectrum needs to be specified as an energy density spectrum instead of a power density spectrum, and in step 5, it is sound exposures that are summed across the bands to obtain the overall sound exposure, which is then converted to a sound exposure level.

As an example, the modeled received sound exposure levels due to a single 3.3-l (200-cui) airgun are plotted as a function of range and depth in Fig. 6.17. The airgun (i.e., a cylindrical tube filled with compressed air, which is suddenly released into the water) is located at the geographical location that was used for the fin whale example, but at a depth of 6 m, which is typical of seismic survey source depths. The scenario is otherwise the same as previously described. The airgun's source waveform was modeled using the *Cagam* airgun array model (Duncan and Gavrilov 2019). The airgun array model also calculated the signal's energy density spectrum, which was then used in step 3 of the broadband modeling procedure outlined above. Once again, *AcTUP* was used to run *RAMGeo* to carry out the propagation modeling, but this time at 1/3 octave band center frequencies from 7.9 Hz to 1 kHz, which took about 5 minutes. A separate *MATLAB* program was written to carry out the post-processing steps and to plot the results.

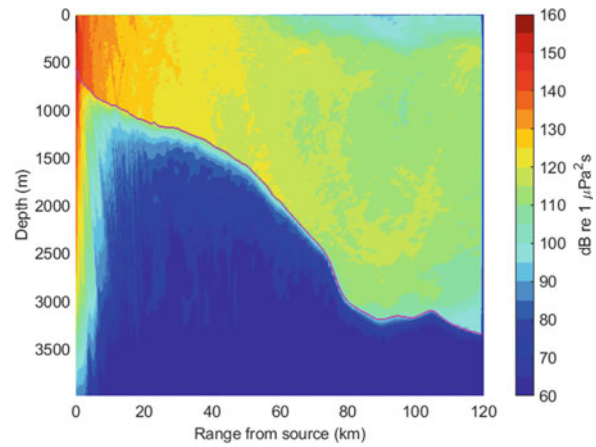



**Fig. 6.17** Received SEL from a 3.3-l (200-cui) airgun at a depth of 6 m as a function of range and depth. The magenta line is the seafloor

Comparing Fig. 6.17 with Fig. 6.16, it can be seen that the broad range of frequencies emitted by the airgun has the effect of smoothing out the fluctuations in the sound field caused by interfering paths. The color scales on these two figures are not directly comparable because Fig. 6.16 gives SPL in dB re 1 μPa whereas Fig. 6.17 presents SEL in dB re 1 $\mu Pa^2 s$. The two are related through:

$$SEL = SPL + 10 \log_{10} T \tag{6.17}$$

where $T$ is the duration of the received signal in seconds, conventionally defined as the duration of the time interval containing 90% of the signal's energy (90% energy signal duration; see Chap. 4).

### 6.5.3 Received Level as a Function of Geographical Position and Depth

The geographical distribution of received sound levels can be modeled by repeating the tonal source modeling procedure (Sect. 6.5.1) or broadband source modeling procedure (Sect. 6.5.2) using bathymetry profiles appropriate for different directions from the source. For long-range modeling, it may also be necessary to make the sound speed profile a function of range and direction. This is called $N \times$ 2D modeling and is adequate in most circumstances, but is less accurate than running a fully 3D propagation model in situations involving sound propagating across steeply sloping seabeds, or in some special situations in which horizontal sound speed gradients become significant.

The result is a 3D grid of the received level as a function of range, depth, and azimuth (i.e., direction in the horizontal plane). To create a 2D map of the sound field, it is necessary to extract some measure of the sound field in the vertical dimension and then interpolate that in the horizontal plane, with the appropriate measure depending on the purpose of the modeling. For example, in environmental impact assessments, it is common to use the maximum level at any depth in the water column, or the maximum level in a depth range corresponding to the diving range of an animal of interest.

Here we illustrate $N \times$ 2D modeling using the previous two examples, but this time carrying out the propagation modeling with bathymetry appropriate for each of the 37 tracks shown in Fig. 6.18. These were set at 10° increments in azimuth, with some adjustment and an extra track inserted in the inshore direction to improve the definition of the received field in the vicinity of the two capes. *MATLAB* programs were written to automate the various steps of the process.

Results are plotted in Fig. 6.19 for the fin whale and the airgun. In both cases, the plots are of the maximum received level over depth, but once again, they are not directly comparable because SPL was plotted for the fin whale, whereas SEL was plotted for the airgun.

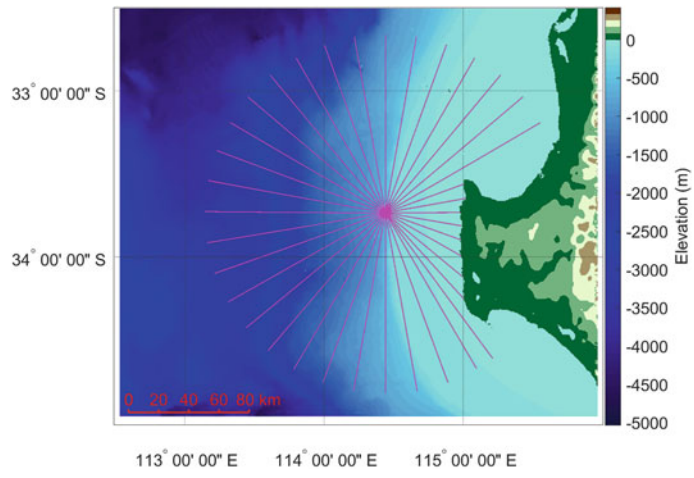



**Fig. 6.18** Map showing the bathymetry off the southwest coast of Australia. The lines radiating from the chosen source location show the tracks along which propagation was modeled

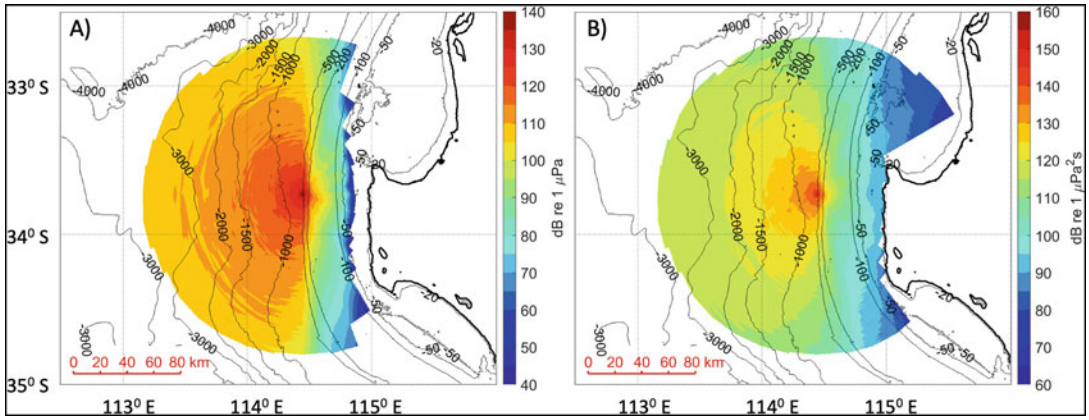



**Fig. 6.19** (**a**) Map of maximum SPL over depth as a function of geographical position due to a fin whale calling at a depth of 50 m off the southwest coast of Australia. (**b**) Map of maximum SEL over depth due to a single firing of an airgun of volume 3.3 l (200 cui) at a depth of 6 m

### 6.5.4 Received Level as a Function of Geographical Position and Depth for a Directional Source

Another level of complexity occurs when the source emits sound differently in different directions. We illustrate this for an airgun array typical of those used for offshore seismic surveys. In this case, the array consists of 30 individual airguns of different sizes arranged in a 21-m wide by 15-m long rectangular array, with all airguns at the same depth of 6 m. The total volume of the compressed air released when the airguns fire is 55.7 l (3400 cui), and the tow direction is towards the North. The *Cagam* airgun array model was used to calculate a representative source spectrum corresponding to the direction of each of the propagation tracks shown in Fig. 6.18. Apart from using a different source spectrum for each direction, the procedure for calculating the received levels was identical to that described in the previous section for the single airgun.

The maximum received SEL at any depth is plotted in Fig. 6.20a, which uses the same color scale as Fig. 6.19b. The array produced higher levels overall, and the sound field was more directional, with distinct maxima east, west, and to a lesser extent, north and south from the source. Figure 6.20b combines range-depth plots for the 90° and 270° azimuths in a single plot, which illustrates the contrasting sound attenuation rates in the upslope and downslope directions.

### 6.5.5 Modeling Limitations and Practicalities

Provided the chosen propagation modeling approach is appropriate for the task, the largest uncertainties in the results are likely due to a lack of information on the environment, which includes the bathymetry, seabed composition, and water column sound speed profile. Bathymetry and water column sound speed profiles are often straightforward to measure or can be obtained from databases, but knowledge of the acoustic properties of the seabed is often poor (i.e., unavailable, patchy, and uncertain) and the parameters that contribute to the geoacoustics (e.g., sediment composition, density, and thickness) vary over space and not coherently (Erbe et al. 2021). Moreover, seabed properties tens or even hundreds of meters below the seafloor may be important when modeling low-frequency propagation (Etter 2018). As a result, it is often prudent to carry out modeling with several

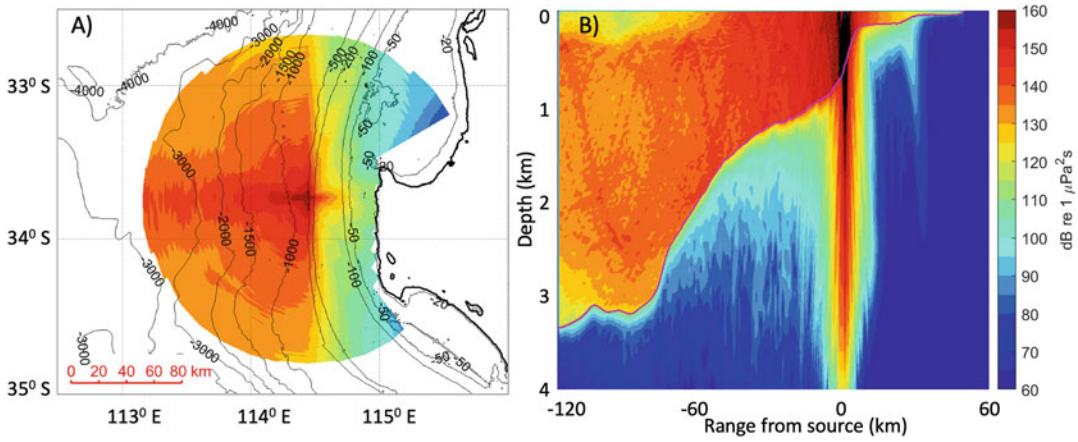



**Fig. 6.20** (**a**) Map of maximum SEL over depth as a function of geographical position due to a single firing of a typical airgun array off the southwest coast of Australia. The total volume of the airguns in the array was 55.7 l (3400 cui), and the array was at a depth of 6 m. The tow direction of the array was northwards. (**b**) Received SEL from the same airgun array as a function of range and depth. The source was at 0-km range, negative ranges correspond to the 270° azimuth (i.e., west of the source) and positive ranges correspond to the 90° azimuth (i.e., east of the source). The magenta line is the seafloor. Colorbar applies to both panels

different sets of seabed properties in order to obtain an estimate of the uncertainty in the results.

The use of $N \times$ 2D rather than fully 3D modeling in the above examples may introduce some inaccuracies for cross-slope propagation paths, which in this case are to the north and south of the source. The effect of the sloping bathymetry would be to deflect the sound towards the downslope direction, slightly increasing levels downslope and decreasing them upslope.

The modeling methods described above treat the source as an ideal point source, which is a good approximation provided the receiver is much farther away from the source than the dimensions of the source. Modeling received levels close to a large source such as an airgun array requires a different and more computationally intensive approach in which the individual airguns in the array are treated as separate sources, and their signals are combined, taking account of their relative phases at the receiver locations. The same approach accounts for the full 3D directivity of the source, rather than just the horizontal directivity, as was the case for the example in Sect. 6.5.4. Combining this approach with a process called Fourier synthesis (Jensen et al. 2011) allows the received waveforms to be simulated, which allows other signal measures such as peak sound pressure levels ($SPL_{pk}$) to be calculated. Calculating $SPL_{pk}$ by this means works well at short ranges but tends to overestimate levels at longer ranges because the propagation models do not properly account for seabed and sea surface scattering effects that broaden the peaks and reduce their amplitudes.

Simple propagation modeling tasks such as those described in Sects. 6.5.1 and 6.5.2 can be carried out using free propagation modeling tools such as the *Acoustics Toolbox* and *AcTUP*, with the addition of some relatively straightforward post-processing coded in any convenient programming language. However, when $N \times$ 2D modeling in multiple directions is required, it becomes desirable to automate the process of interpolating bathymetry profiles from databases, generating sound speed profile files, initiating multiple runs of the propagation model, calculating received levels, interpolating and plotting results, etc. Most organizations that routinely carry out this type of modeling have written their own proprietary software for these tasks. To the authors' knowledge, there is no freely available software package with all of these capabilities, although there is at least one commercially available package.

## 6.6 Summary

Sound propagation under water is a complex process. Sound does not propagate along straight-line transmission paths. Rather, it reflects, refracts, and diffracts. It scatters off rough surfaces (such as the sea surface and the seafloor) and off reflectors within the water column (e.g., gas bubbles, fish swim bladders, and suspended particles). It is transmitted into the seafloor and partially lost from the water. It is converted into heat by exciting molecular vibrations. There are common misconceptions about sound propagation in water, such as "low-frequency sound does not propagate in shallow water," "over hard seafloors, all sound is reflected, leading to cylindrical spreading," and "over soft seafloors, sound propagates spherically." This chapter aimed to remove common misconceptions and empower the reader to comprehend sound propagation phenomena in a range of environments and appreciate the limitations of widely used sound propagation models. The chapter began by deriving the sonar equation for a number of scenarios including animal acoustic communication, communication masking by noise, and acoustic surveying of animals. It introduced the concept of the layered ocean, presenting temperature, salinity, and resulting sound speed profiles. These were needed to develop the most common concepts of sound propagation under water: ray tracing and normal modes. The chapter computed Snell's law, reflection and transmission coefficients, and Lloyd's mirror. It provided an overview of publicly available sound propagation software (including wavenumber integration and parabolic equation models). It concluded with a few practical examples of modeling propagation loss for whale song and a seismic airgun array.

## 6.7 Additional Resources

- Dan Russell's Acoustics and Vibration Animations: https://www.acs.psu.edu/drussell/demos.html
- The Discovery of Sound in the Sea (DOSITS; https://dosits.org/) website has over 400 pages of content in three major sections including the science of underwater sound and how people and marine animals use underwater sound to conduct activities for which light is used in air. The website has been the foundational resource of the DOSITS Project, providing information at a beginner and advanced level, based on peer-reviewed science (Vigness-Raposa et al. 2016, 2019). The web structure has been transformed into structured tutorials that provide a streamlined, progressive development of knowledge. The tutorial layout allows a user to proceed from one topic to the next in sequence or jump to a specific topic of interest. The three tutorials focus on the science of underwater sound, the potential effects of underwater sound on marine animals, and the ecological risk assessment process for determining possible effects from a specific sound source. Additional resources have been developed to provide the underwater acoustics content in different formats, including instructional videos and webinars. Finally, there are print publications (an educational booklet and a trifold brochure) available in hard copy or PDF format and two eBooks available for free on the iBooks Store, including Book I: Importance of Sound in the Sea and Book II: Science of Underwater Sound.

## References

Ainslie MA (2005) Effect of wind-generated bubbles on fixed range acoustic attenuation in shallow water at 1–4 kHz. J Acoust Soc Am 118(6):3513–3523. https://doi.org/10.1121/1.2114527

Ainslie MA, McColm JG (1998) A simplified formula for viscous and chemical absorption in sea water. J Acoust Soc Am 103(3):1671–1672. https://doi.org/10.1121/1.421258

American National Standards Institute (2015) Bioacoustical Terminology (ANSI S3.20-2015, R 2020). Acoustical Society of America, New York

Au WWL, Moore PWB (1984) Receiving beam patterns and directivity indices of the Atlantic bottlenose dolphin *Tursiops truncatus*. J Acoust Soc Am 75(1):255–262. https://doi.org/10.1121/1.390403

Benoit-Bird KJ, Waluk CM (2020) Exploring the promise of broadband fisheries echosounders for species discrimination with quantitative assessment of data processing effects. J Acoust Soc Am 147(1):411–427. https://doi.org/10.1121/10.0000594

Benoit-Bird KJ, Moline MA, Southall BL (2017) Prey in oceanic sound scattering layers organize to get a little help from their friends. Limnol Oceanogr 62(6): 2788–2798. https://doi.org/10.1002/lno.10606

Collins MD (1993) An energy-conserving parabolic equation for elastic media. J Acoust Soc Am 94(2):975–982

Collis JM, Siegmann WL, Jensen FB, Zampolli M, Küsel ET, Collins MD (2008) Parabolic equation solution of seismo-acoustics problems involving variations in bathymetry and sediment thickness. J Acoust Soc Am 123(1):51–55. https://doi.org/10.1121/1.2799932

Duncan AJ, Gavrilov AN (2019) The CMST Airgun Array model—a simple approach to modeling the underwater sound output from seismic Airgun arrays. IEEE J Ocean Eng 44(3):589–597. https://doi.org/10.1109/JOE.2019.2899134

Erbe C (2000) Detection of whale calls in noise: performance comparison between a beluga whale, human listeners and a neural network. J Acoust Soc Am 108(1):297–303. https://doi.org/10.1121/1.429465

Erbe C (2002) Underwater noise of whale-watching boats and its effects on killer whales (*Orcinus orca*). Mar Mamm Sci 18(2):394–418. https://doi.org/10.1111/j.1748-7692.2002.tb01045.x

Erbe C (2008) Critical ratios of beluga whales (*Delphinapterus leucas*) and masked signal duration. J Acoust Soc Am 124(4):2216–2223. https://doi.org/10.1121/1.2970094

Erbe C (2015) The maskogram: a tool to illustrate zones of masking. Aquat Mamm 41(4):434–443. https://doi.org/10.1578/AM.41.4.2015.434

Erbe C, Farmer DM (1998) Masked hearing thresholds of a beluga whale (*Delphinapterus leucas*) in icebreaker noise. Deep Sea Res II 45(7):1373–1388. https://doi.org/10.1016/S0967-0645(98)00027-7

Erbe C, Liong S, Koessler MW, Duncan AJ, Gourlay T (2016a) Underwater sound of rigid-hulled inflatable boats. J Acoust Soc Am 139(6):EL223–EL227. https://doi.org/10.1121/1.4954411

Erbe C, Parsons M, Duncan AJ, Allen K (2016b) Underwater acoustic signatures of recreational swimmers, divers, surfers and kayakers. Acoust Aust 44(2): 333–341. https://doi.org/10.1007/s40857-016-0062-7

Erbe C, Reichmuth C, Cunningham KC, Lucke K, Dooling RJ (2016c) Communication masking in marine mammals: a review and research strategy. Mar Pollut Bull 103:15–38. https://doi.org/10.1016/j.marpolbul.2015.12.007

Erbe C, Dunlop R, Jenner KCS, Jenner M-NM, McCauley RD, Parnum I, Parsons M, Rogers T, Salgado-Kent C (2017a) Review of underwater and in-air sounds emitted by Australian and Antarctic marine mammals. Acoust Aust 45:179–241. https://doi.org/10.1007/s40857-017-0101-z

Erbe C, Parsons M, Duncan AJ, Osterrieder S, Allen K (2017b) Aerial and underwater sound of unmanned aerial vehicles (UAV, drones). J Unmanned Veh Syst 5(3):92–101. https://doi.org/10.1139/juvs-2016-0018

Erbe C, Williams R, Parsons M, Parsons SK, Hendrawan IG, Dewantama IMI (2018) Underwater noise from airplanes: an overlooked source of ocean noise. Mar Pollut Bull 137:656–661. https://doi.org/10.1016/j.marpolbul.2018.10.064

Erbe C, Peel D, Smith JN, Schoeman RP (2021) Marine acoustic zones of Australia. J Mar Sci Eng 9(3):340. https://doi.org/10.3390/jmse9030340

Etter PC (2018) Underwater acoustic modeling and simulation, 5th edn. CRC Press, Boca Raton, FL. https://doi.org/10.1201/9781315166346

Fletcher H (1940) Auditory patterns. Rev Mod Phys 12: 47–65

François RE, Garrison GR (1982a) Sound absorption based on ocean measurements: part I: pure water and magnesium sulphate contributions. J Acoust Soc Am 72(3):896–907

François RE, Garrison GR (1982b) Sound absorption based on ocean measurements: part II: boric acid contribution and equation for total absorption. J Acoust Soc Am 72(6):1879–1890

Godin OA (2008) Sound transmission through water–air interfaces: new insights into an old problem. Contemp Phys 49(2):105–123. https://doi.org/10.1080/00107510802090415

Goh JT, Schmidt H (1996) A hybrid coupled wavenumber integration approach to range-dependent seismoacoustic modeling. J Acoust Soc Am 100(3): 1409–1420. https://doi.org/10.1121/1.415988

Hall MV (1989) A comprehensive model of wind-generated bubbles in the ocean and predictions of the effects on sound propagation at frequencies up to 40 kHz. J Acoust Soc Am 86(3):1103–1117. https://doi.org/10.1121/1.398102

Hall MV (2004) Preliminary analysis of the applicability of adiabatic modes to inverting synthetic acoustic data in shallow water over a sloping sea floor. IEEE J Ocean Eng 29(1):51–58. https://doi.org/10.1109/JOE.2003.823315

Jensen FB, Kuperman WA, Porter MB, Schmidt H (2011) Computational Ocean acoustics, 2nd edn. Springer, New York

Kloser RJ, Macaulay GJ, Ryan TE, Lewis M (2013) Identification and target strength of orange roughy (Hoplostethus atlanticus) measured in situ. J Acoust Soc Am 134(1):97–108

Koessler MW (2016) Modelling of underwater acoustic propagation over elastic, range-dependent seabeds. Ph. D. Thesis, Curtin University, Perth, WA, Australia

Kuehne LM, Erbe C, Ashe E, Bogaard LT, Collins MS, Williams R (2020) Above and below: military aircraft noise in air and under water at Whidbey Island, Washington. J Mar Sci Eng 8(11):923. https://doi.org/10.3390/jmse8110923

Locarnini RA, Mishonov AV, Baranova OK, Boyer TP, Zweng MM, Garcia HE, Reagan JR, Seidov D, Weathers K, Paver CR, Smolyar I (2018) World Ocean atlas 2018, volume 1: temperature. National Oceanic and Atmospheric Administration

Mackenzie KV (1981) Nine-term equation for sound speed in the oceans. J Acoust Soc Am 70:807–812

Medwin H, Clay CS (1998) Chapter 2 - sound propagation. In: Medwin H, Clay CS (eds) Fundamentals of acoustical oceanography. Academic Press, San Diego, pp 17–69. https://doi.org/10.1016/B978-012487570-8/50004-0

Moore BCJ (2013) An introduction to the psychology of hearing. Brill, Leiden, The Netherlands

Parsons MJG, Parnum IM, Allen K, McCauley RD, Erbe C (2014) Detection of sharks with the Gemini imaging sonar. Acoust Australia 42(3):185–189

Parsons MJG, Duncan AJ, Parsons SK, Erbe C (2020) Reducing vessel noise: an example of a solar-electric passenger ferry. J Acoust Soc Am 147(5):3575–3583. https://doi.org/10.1121/10.0001264

Porter MB (1990) The time-marched fast-field program (FFP) for modeling acoustic pulse propagation. J Acoust Soc Am 87(5):2013–2023. https://doi.org/10.1121/1.399329

Porter MB, Bucker HP (1987) Gaussian beam tracing for computing ocean acoustic fields. J Acoust Soc Am 82(4):1349–1359. https://doi.org/10.1121/1.395269

Porter M, Reiss EL (1984) A numerical method for ocean-acoustic normal modes. J Acoust Soc Am 76(1):244–252. https://doi.org/10.1121/1.391101

Schmidt H, Glattetre J (1985) A fast field model for three-dimensional wave propagation in stratified environments based on the global matrix method. J Acoust Soc Am 78(6):2105–2114. https://doi.org/10.1121/1.392670

Sirovic A, Hildebrand JA, Wiggins SM (2007) Blue and fin whale call source levels and propagation range in the Southern Ocean. J Acoust Soc Am 122(2):1208–1215. https://doi.org/10.1121/1.2749452

Urick RJ (1983) Principles of underwater sound, 3rd edn. McGraw Hill, New York

Vigness-Raposa KJ, Scowcroft G, Morin H, Knowlton C, Miller JH, Ketten DR, Popper AN (2016) Discovery of sound in the sea: resources for decision makers. Proc Meet Acoust 27(1):010008. https://doi.org/10.1121/2.0000257

Vigness-Raposa KJ, Scowcroft G, Morin H, Knowlton C, Miller JH, Ketten DR, Popper AN (2019) Discovery of sound in the sea: communicating underwater acoustics research to decision makers. Proc Meet Acoust 37(1):025001. https://doi.org/10.1121/2.0001204

Westwood EK, Tindle CT, Chapman NR (1996) A normal mode model for acousto-elastic ocean environments. J Acoust Soc Am 100(6):3631–3645. https://doi.org/10.1121/1.417226

Zweng MM, Reagan JR, Seidov D, Boyer TP, Locarnini RA, Garcia HE, Mishonov AV, Baranova OK, Weathers K, Paver CR, Smolyar I (2018) World Ocean atlas 2018, volume 2: salinity. National Oceanic and Atmospheric Administration

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



# Analysis of Soundscapes as an Ecological Tool

7

Renée P. Schoeman, Christine Erbe, Gianni Pavan, Roberta Righini, and Jeanette A. Thomas

## 7.1 Introduction

Whether listening in a forest or on an open plain, by the side of a river or in the ocean, at the outskirts of suburbia or right downtown, the Earth abounds with sounds. The use of the term "soundscape" in the literature has increased rapidly since 2000 (Fig. 7.1) and can be traced back to Southworth's (1969) article on the sonic environment of Boston, MA, USA. The Canadian music composer and researcher Schafer later defined soundscapes as "the auditory properties of landscapes" (Schafer 1977). Schafer was a pioneer in highlighting the need for soundscape research and management. In his book, *The New Soundscape*, Schafer and his students documented rapid changes in soundscapes over the course of human civilization (Schafer 1969). Common settings of primitive cultures surrounded by an abundance of natural sounds (i.e., wind, water, animals, etc.) rapidly changed after the Industrial Revolution to cities dominated by sounds from machinery. Schafer further noticed that most people had ceased to listen to the sounds of the environment and actively tried to ignore unpleasant sound (i.e., noise). With the goals of studying and archiving soundscapes, creating public awareness of noise pollution, and creating healthy soundscapes through acoustic design, Schafer founded the World Soundscape Project (WSP 1972–1979; Torigoe 1982). Soundscape studies by the WSP were human-centered, focusing on the acoustic composition of cities and villages, studying only humans as receivers of acoustic information, and emphasizing the negative effects of noise on humans (Truax 1984, 1996). Krause (1987, 1993) adopted an animal-centered approach to the study of soundscapes. He recorded and archived sounds of different animal species as well as of entire ecosystems. According to Krause, acoustic sampling of an area over a period of time and under different conditions allows us to study, and ultimately predict, how human-induced changes might affect ecosystems (Krause 1987).

While the term "soundscape" has different uses in the literature, the International Organization for Standardization officially defined "soundscape" as "an acoustic environment as perceived or experienced and/or understood by a person or people, in context" and "acoustic environment" as the "sound at the receiver from all

Jeanette A. Thomas (deceased) contributed to this chapter while at the Department of Biological Sciences, Western Illinois University-Quad Cities, Moline, IL, USA

R. P. Schoeman (✉) · C. Erbe
Centre for Marine Science and Technology, Curtin University, Perth, WA, Australia
e-mail: renee.koper@postgrad.curtin.edu.au; c.erbe@curtin.edu.au

G. Pavan · R. Righini
Centro Interdisciplinare di Bioacustica e Ricerche Ambientali, University of Pavia, Pavia, Italy
e-mail: gianni.pavan@unipv.it

© The Author(s) 2022

C. Erbe, J. A. Thomas (eds.), *Exploring Animal Behavior Through Sound: Volume 1*,
https://doi.org/10.1007/978-3-030-97540-1_7

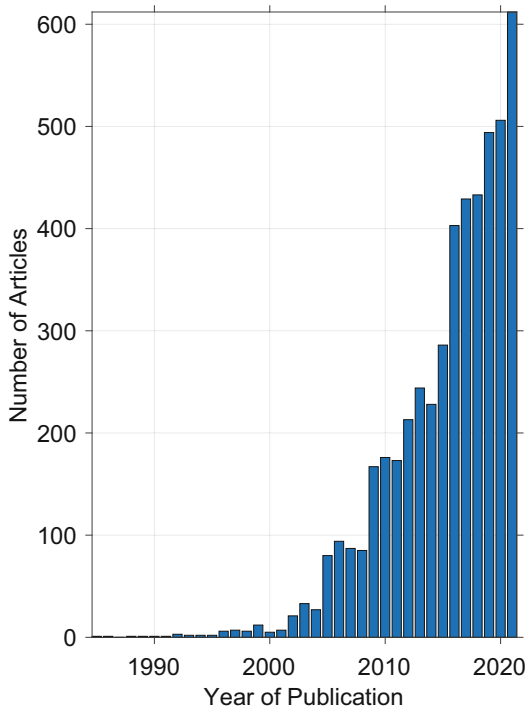



**Fig. 7.1** Number of articles with "soundscape" in the abstract, listed by Scopus, versus publication year; retrieved 10 June 2022

sound sources as modified by the environment" (International Organization for Standardization [ISO] 2014). A soundscape is thus a perceptual construct that requires a human listener, while the acoustic environment is a physical phenomenon, extending in frequency beyond the human hearing limits, including infrasounds and ultrasounds. In the field of underwater acoustics, however, a soundscape is the "characterization of the ambient sound in terms of its spatial, temporal and frequency attributes, and the types of sources contributing to the sound field" (International Organization for Standardization [ISO] 2017). "Soundscape" in underwater acoustics thus does not require a listener. In essence, the usage of the term "soundscape" in the literature is variable and perhaps related to specific research objectives (Scarpelli et al. 2020).

The components of a soundscape may be grouped by their origin. Sounds produced by animals are grouped as biophony, sounds produced by atmospheric or geophysical events make up the geophony, and sounds produced by human activities or machinery are referred to as anthropophony (Fig. 7.2; Krause 2008). Sounds created by machinery (including power generators, motors, etc.) are sometimes grouped as technophony (Mullet et al. 2016), which is the component of anthropophony typically associated with noise pollution. The identification of soundscape components is a key element in the research field of ecoacoustics, which investigates the relationship of natural and anthropogenic sounds with the environment on a range of scales in space and time (Farina and Gage 2017). The research field of soundscape ecology investigates the interaction of organisms with their environment, mediated through sound (Pijanowski et al. 2011a, b). For example, sound sources distributed within an environment provide acoustic cues (i.e., soundmarks), by which animals can orientate, navigate, and make habitat choices (Slabbekoorn and Bouton 2008). Under the Acoustic Habitat Hypothesis, the habitats that sound-dependent species select and occupy exhibit acoustic characteristics that suit a species' functional needs and match its sound production and reception capabilities (Mullet et al. 2017a). Acoustic habitat specialists are species whose acoustic habitat is unique and vital to its functional needs, while acoustic habitat generalists occupy acoustic habitats that are less than unique but still important to the species' functional needs (Mullet et al. 2017a). Under the Acoustic Adaptation Hypothesis, the sounds of soniferous animals evolved to optimize propagation within the animals' habitat (Morton 1975), characterized by its soundscape and sound propagation conditions. Under the Acoustic Niche Hypothesis, animals evolved species-specific sounds in certain frequency bands and temporal patterns to minimize competition (i.e., masking) with sounds from other animals and the environment (Krause 1993). An interesting and related question is how animal (and human) listeners make sense of the myriad of sounds received from all directions, overlapping in frequency and time, and thus masking each other. A listener must separate the parts belonging to different sources and merge the parts belonging to the same source to make sense

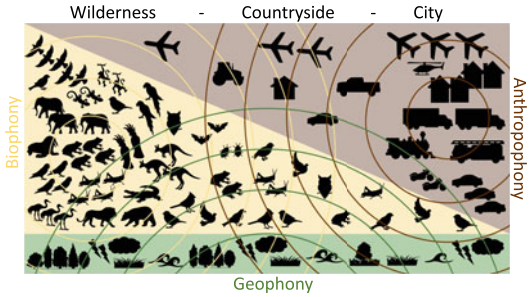



**Fig. 7.2** Sketch of the sound sources within soundscapes ranging from wilderness to countryside, to city. Biophony decreases and anthropophony increases while the geophony might vary comparatively little. Example species are sketched along the way with decreasing density and biodiversity. Acoustic habitat generalists occur in multiple, different soundscapes, while acoustic habitat specialists only occur in quite specific soundscapes (Mullet et al. 2017a)

of the acoustic scene. This is called auditory scene analysis (Bregman 1990; Lewicki et al. 2014).

Natural soundscapes are appreciated for their esthetic and recreational value (e.g., Davies et al. 2013; Francis et al. 2017; Franco et al. 2017) and also have a significant ecological and scientific value. Soundscapes should, therefore, be considered a natural resource, worthy of study, management, and conservation (National Park Service [NPS] 2000; Farina and Gage 2017; Pavan 2017). How many undisturbed soundscapes remain in this world of decreasing biodiversity, changes in land-use, and rising anthropogenic noise? Can the soundscape of a pristine habitat function as a model to restore a degraded habitat (Pavan 2017; Gordon et al. 2019; Righini and Pavan 2020)? This chapter gives an overview of terrestrial and aquatic soundscapes, outlines how soundscapes may change or have changed over time, provides tools for analyzing and quantifying soundscapes, and discusses how passive acoustic monitoring applies to soundscape ecology research, management, and conservation.

## 7.2 Terrestrial Soundscapes

Terrestrial soundscapes may vary widely within as well as between ecosystems (e.g., Krause 2012; Yip et al. 2017; Priyadarshani et al. 2018). While some soundscapes might have been studied more than others (Scarpelli et al. 2020), there often are key sounds (i.e., sounds characteristic for an ecosystem) by which an ecosystem may be identified. For example, a listener may identify the terrestrial soundscape of a nearshore ecosystem off central California, USA, by the barks of California sea lions (*Zalophus californianus*), the squawks of sea gulls (*Larus californicus*), and the tapping sounds made by sea otters (*Enhydra lutris*) that use a rock to crack-open shellfish.

### 7.2.1 Biophony

The terrestrial biophony includes sounds produced by insects (e.g., Brady 1974; Römer and Lewald 1992; Polidori et al. 2013), anurans (e.g., Cunnington and Fahrig 2010; Zhang et al. 2017), reptiles (e.g., Crowley and Pietruszka 1983; Galeotti et al. 2005), birds (e.g., Lengagne et al. 1999; Charrier et al. 2001; Catchpole and Slater 2008), bats (e.g., Gadziola et al. 2012; Prat et al. 2016), and other mammals (such as dogs and seals; e.g., van Opzeeland et al. 2010; Mumm and Knörnschild 2014; Bowling et al. 2017). Typically, multiple (vocal) taxa occur in the same environment and so, evidence for the Acoustic Niche Hypothesis has been demonstrated in various ecosystems among insects (Sueur 2002), anurans (Villanueva-Rivera 2014), birds (Azar and Bell 2016), and a combination of species (Hart et al. 2015).

Terrestrial soundscape ecology studies have been dominated by research on birds (Ferreira et al. 2018). Most bird species are diurnal vocalizers, with peak activity at dawn and dusk. Birds may emit single calls as well as sounds arranged into long and complex songs (Fig. 7.3). Calls have a variety of functions and are, for example, produced to raise alarm (Gill and Bierema 2013), contact conspecifics (Bond and Diamond 2005), or beg for food (Klenova 2015). While bird song was long thought to be an exclusive male trait used for territorial defense and female attraction, there is mounting evidence

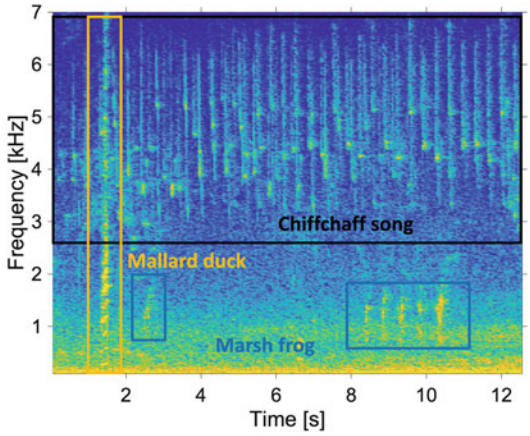



**Fig. 7.3** Soundscape of a temperate forest at dusk showing song of the chiffchaff (*Phylloscopus collybita*), squawks of a mallard duck (*Anas platyrhynchos*), and calls from a marsh frog (*Pelophylax ridibundus*)

that female bird song is globally widespread and used for territorial and reproductive purposes (Odom et al. 2014). Terrestrial birds primarily communicate within the frequency range of human hearing, with recorded fundamental frequencies (see Chap. 4) as low as 23 Hz for southern cassowary (*Casuarius casuarius*; Mack and Jones 2003) and as high as 13 kHz for the Ecuadorian hillstar hummingbird (*Oreotrochilus chimborazo*; Duque et al. 2018). Marine birds that are heard within terrestrial soundscapes produce calls with fundamental frequencies <2 kHz (e.g., Charrier et al. 2001; Bourgeois et al. 2007; Cure et al. 2009; Mulard et al. 2009; Dentressangle et al. 2012). Lesser-known sounds of birds are those produced by wings while in flight and while perched (Clark 2021). Because these sounds may be audible to the animal itself, conspecifics, and other species (e.g., predators and prey), Clark (2021) suggested that these sounds may be selected to evolve from by-product to communication signal.

Insects are another common source of biophony, with seasonal and diurnal choruses produced by cicadas and crickets at dominant frequencies between 2 and 50 kHz (Bennet-Clark 1970; Robillard et al. 2013; Hart et al. 2015; Buzzetti et al. 2020). These typically male insect choruses, produced to attract females, can be intense and potentially affect the timing and frequency of other species' vocalizations. Hart et al. (2015), for example, found that birds in a Costa Rican tropical rainforest either ceased vocalizing or changed their call frequency to avoid acoustic overlap with cicada choruses (Fig. 7.4). As do birds, insects produce sounds in flight, with dominant frequencies between 140 and 250 Hz (Fig. 7.5; Kawakita and Ichikawa 2019).

Social wasps, honeybees, bumble bees, and some hoverflies produce sounds with dominant frequencies between 152 and 317 Hz when attacked by predators, potentially as a warning signal (Rashed et al. 2009). Smaller velvet ants (family of wasps) also produce distress calls but at higher frequencies between 4 and 17 kHz (Polidori et al. 2013). Ants produce distress calls extending in frequency above 70 kHz (Pavan et al. 1997).

In many anuran species, males aggregate and produce evening choruses of varying complexity to advertise for females (i.e., courtship vocalizations; Grafe 2005). Most male anuran species cycle air through a vocal sac to produce calls with main energy between 400 Hz and 10 kHz (Fig. 7.5c; Cunnington and Fahrig 2010; Narins and Meenderink 2014; Villanueva-Rivera 2014), although some species produce sounds that extend into the ultrasonic range (i.e., >20 kHz; Feng et al. 2006; Arch et al. 2008). White-lipped frogs (*Leptodactylus albilabris*) also thump their vocal sac on the underlying substrate while vocalizing, thereby creating a seismic signal, which potentially plays a role in seismic communication with conspecifics (Narins 1990).

Courtship vocalizations have also been recorded for at least 35 species of tortoises. Call characteristics of 11 tortoise species were studied in detail by Galeotti et al. (2005), revealing dominant frequencies between 110 and 600 Hz and energy between 100 Hz and 3 kHz. Snakes may produce a broadband hiss (3–13 kHz; Young 1991), rattle (2–23 kHz; Young and Brown 1993), or rasping sound (200 Hz–11 kHz; Young 2003) when threatened. Crocodiles produce sounds with main energy <2 kHz (e.g.,

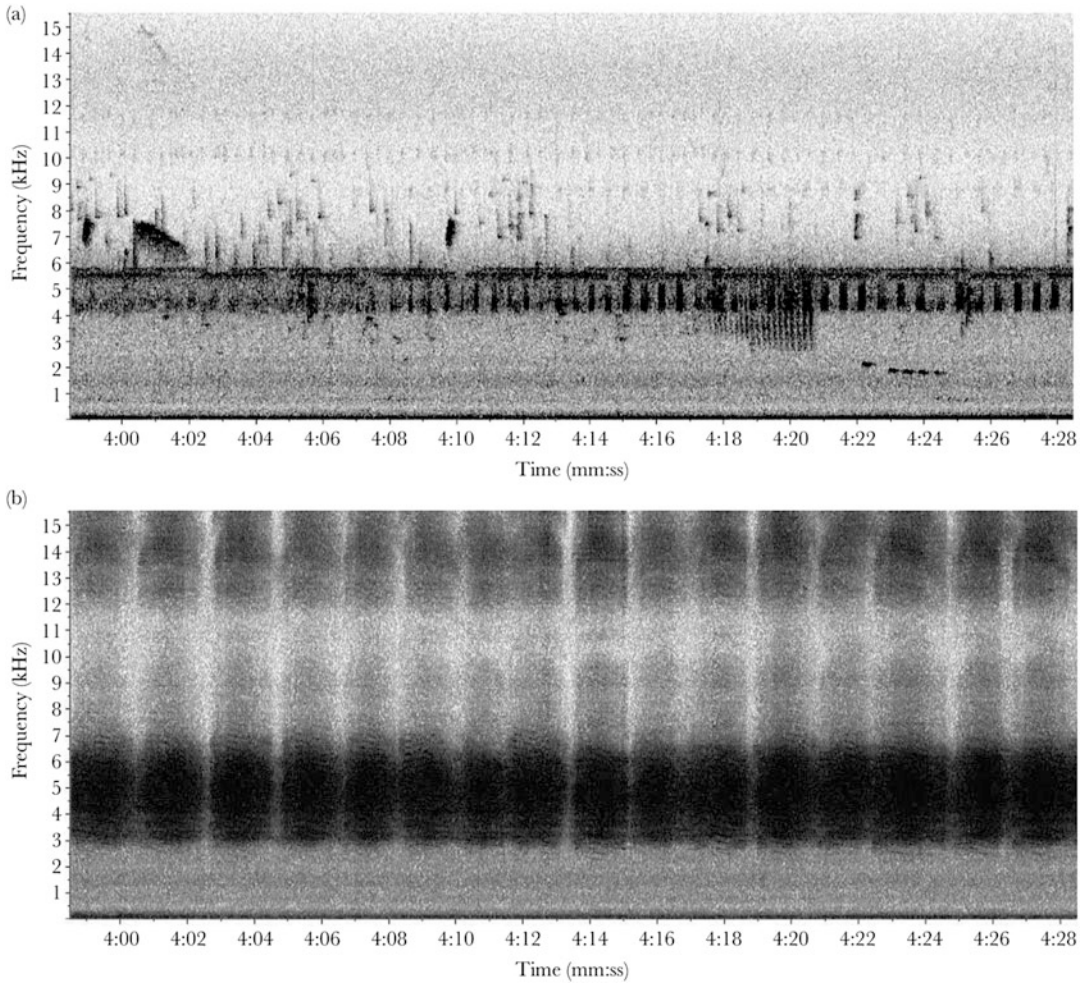



**Fig. 7.4** A comparison of the soundscapes at two different moments of the morning in a secondary wet forest at Las Cruces Biological Station, Costa Rica. Top spectrogram recorded minutes prior to the onset of *Zammara smaragdina* cicada morning choruses, displaying vocalizations from seven bird species (*Arremon aurantiirostris, Picumnus olivaceus, Arremon torquatus, Catharus aurantiirostris, Arremon aurantiirostris, Phaeothlypis fulvicauda,* and *Formicarius analis*). Bottom spectrogram recorded at the same location just after the onset of cicada morning choruses. © Hart et al. (2015); https://academic.oup.com/view-large/figure/79529274/beheco_arv018_f0001.jpeg. Published under CC BY 3.0; https://creativecommons.org/licenses/by/3.0/

Vergne et al. 2009, 2011; Reber et al. 2017). Crocodile hatchlings emit calls before, during, and after hatching, which function to synchronize hatching, alert the mother to their due arrival, and stay in contact (Vergne et al. 2011; Chabert et al. 2015). Adult crocodiles produce calls during courtship, during territorial defense, and to maintain group cohesion with offspring (Fig. 7.6; Vergne et al. 2009; Reber et al. 2017).

Mammalian species vocalize at frequencies that, for some taxa, are inversely related to their body size (Bowling et al. 2017). African elephants (*Loxodonta africana*) and Asian elephants (*Elephas maximus*), for example, vocalize within the infrasonic range (i.e., <20 Hz; fundamental frequency as low as 14 Hz). These low-frequency calls function to coordinate movement and to advertise an individual's

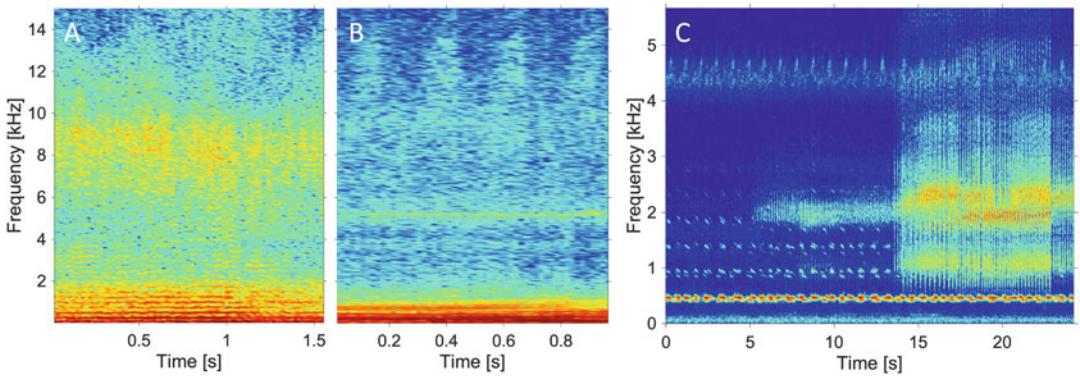



**Fig. 7.5** Spectrograms of the flight sound produced by the European honeybee (*Apis mellifera*; **a**) and the Japanese yellow hornet (*Vespa simillima xanthoptera*; **b**). Sound files from Kawakita and Ichikawa (2019). Spectrogram of chorusing frogs in a pond in Colli Euganei, Italy. Yellow-bellied toad (*Bombina variegata*) with 500-Hz tonals and overtones and the European tree frog (*Hyla arborea*) with higher-pitched, broadband sounds starting at around 5 s and increasing in intensity and bandwidth from 13 s onwards (**c**). Recording courtesy of Marco Pesente

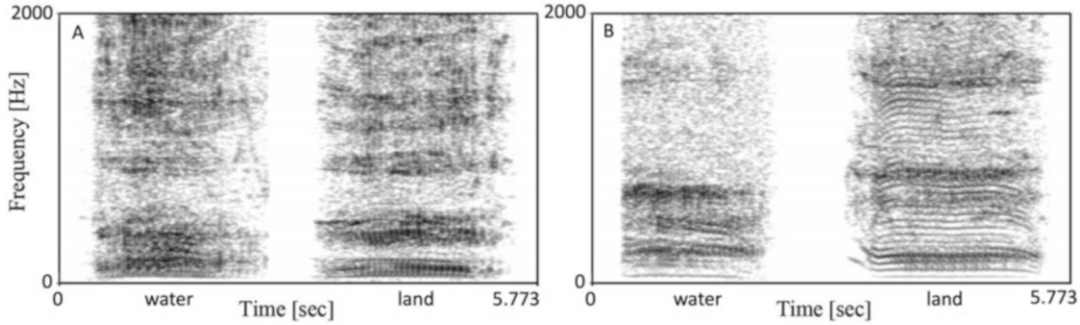



**Fig. 7.6** Male (**a**) and female (**b**) American alligator (*Alligator mississippiensis*) bellows that may be produced during courtship and territorial defense (Vergne et al. 2009). Modified from Reber et al. (2017). © Reber et al. (2017); https://www.nature.com/articles/s41598-017-01948-1/figures/2. Published under CC BY 4.0; https://creativecommons.org/licenses/by/4.0/

reproductive status over distances as far as 2.5 km (Soltis 2010). Elephants also produce vibrations that propagate through the substrate and so provide additional cues to listening conspecifics (Payne et al. 1986; O'Connell-Rodwell et al. 2000). The majority of aerial feeding bats, at the opposite end of the body-size scale, produce short echolocation calls (biosonar) in the ultrasonic range (15–110 kHz), for navigation and hunting (Fenton et al. 1998). Bat social calls, potentially related to agonistic encounters and courtship, are also characterized by harmonics that extend well into the ultrasonic range (Fig. 7.7; Behr and van Helversen 2004; Lattenkamp et al. 2019).

Primate vocalizations cover a wide frequency range from approximately 100 Hz in western gorillas (*Gorilla gorilla;* Salmi et al. 2013) to 16 kHz in pygmy marmosets (*Cebuella pygmaea*; Pola and Snowdon 1975). Primate vocalizations play an important role in intergroup communication, predominantly facilitating social interactions and group movement (Cheney and Seyfarth 1996,

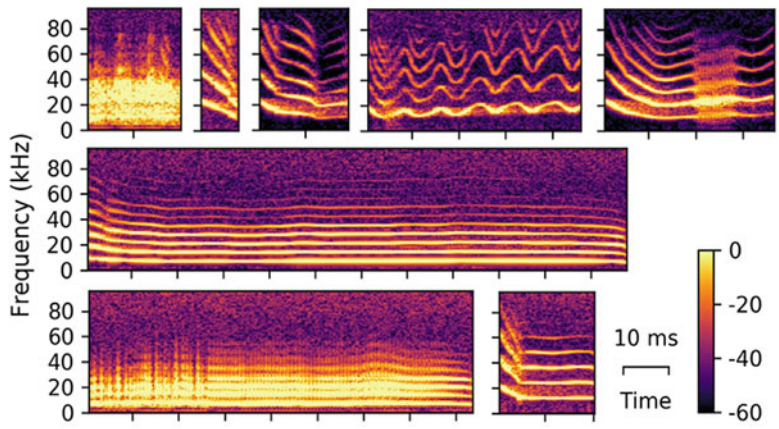



**Fig. 7.7** Common social calls with ultrasonic components emitted by the pale spear-nosed bat (*Phyllostomus discolor*). Modified figure. © Lattenkamp et al. (2019); https://www.frontiersin.org/files/Articles/447704/fevo-07-00116-HTML/image_m/fevo-07-00116-g002.jpg. Published under CC BY 4.0; https://creativecommons.org/licenses/by/4.0/

2018). Primates are also known to use various alarm calls, which were previously suggested to be functionally referential signals (e.g., Cheney and Seyfarth 1996). However, recent studies have shown that primates often use general alarm calls and infer meaning from previous experiences or contextual information (Fichtel 2020).

Marine mammals, such as polar bears (*Ursus maritimus*), pinnipeds (i.e., seals, sea lions, and walruses), and sea otters (*Enhydra lutris nereis*) also produce in-air sounds. Nursing female polar bears frequently emit a low-intensity, repetitive, pulsed sound when initiating or continuing body contact with their cub (20 Hz–2 kHz; Wemmer et al. 1976). Pinnipeds produce in-air sounds with main energy <9 kHz (Fig. 7.8). Mother and pup recognize each other by individually unique calls that help them to reunite amidst all other individuals of the colony (Insley et al. 2010), while males produce individually unique calls during agonistic behavior (e.g., Fernández-Juricic et al. 1999; Van Parijs and Kovacs 2002). Female

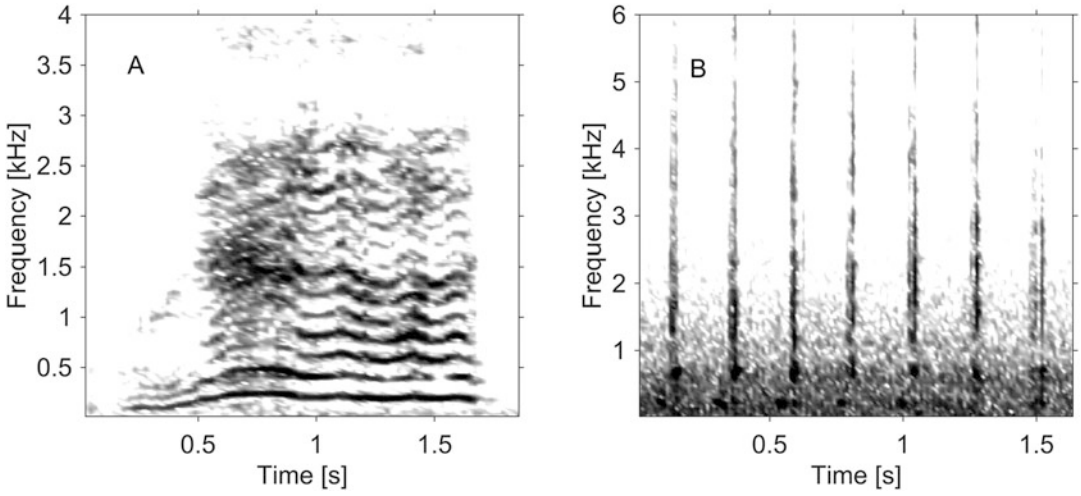



**Fig. 7.8** In-air vocalizations produced by (**a**) a New Zealand fur seal (*Arctocephalus forsteri*) and (**b**) an Australian sea lion (*Neophoca cinerea*). © Erbe et al. (2017); https://doi.org/10.1007/s40857-017-0101-z. Published under CC BY 4.0; https://creativecommons.org/licenses/by/4.0/

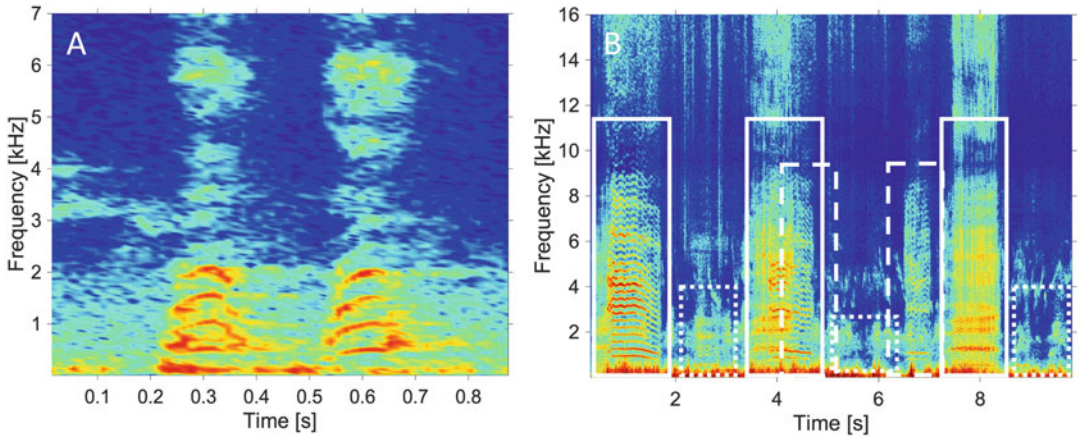



**Fig. 7.9** Example spectrograms of dog barks (**a**) and bleating sheep (**b**). Sheep bleats were produced by an ewe (solid box), her lamb (dashed box), and a distant lamb (dotted box)

and pup sea otters produce individually distinct calls with main energy <5 kHz, which also seem to function as contact calls between separated individuals (McShane et al. 1995).

Urbanized areas may be characterized by the sounds of domesticated animals (i.e., pets and livestock). Dogs bark to greet conspecifics and humans, during play (i.e., excitement), when raising alarm, or when seeking attention (Yin and McCowan 2004), sometimes to the nuisance of the neighborhood (Flint et al. 2014). Barks are short acoustic signals with main energy between 300 Hz and 2.5 kHz (Fig. 7.9), often repeated in bouts (Yin and McCowan 2004). Ewes and their lamb recognize each other by unique calls with main energy <5 kHz (Sèbe et al. 2008), resulting in a cacophony of bleats in lambing season.

### 7.2.2 Geophony

The prevailing geophonic source of sound is wind. Wind acts on vegetation, thereby contributing to sound levels <1 kHz in leafless trees, <4 kHz in leafed trees, and <10 kHz in open grasslands, with a positive correlation between wind speed and sound intensity (Boersma 1997; Bolin 2009). Wind noise may affect the audible range of biological sounds. The detection of bird song in open grasslands in New Zealand significantly decreased with increasing wind speeds from calm (<4 km/h) to windy (>15 km/h) conditions (Priyadarshani et al. 2018). Precipitation also creates sound (Fig. 7.10). Rain increased sound levels within a deciduous forest (Ardennes, France) within the frequency band of 100 Hz to 10 kHz (Lengagne and Slater 2002). The increase in sound levels resulted in a reduction of acoustic communication space (i.e., area over which an individual can communicate with conspecifics) for tawny owls

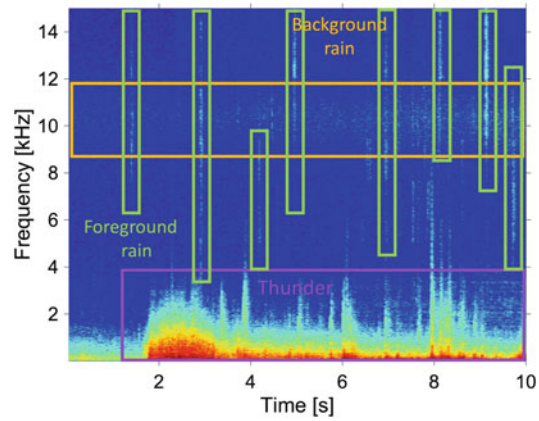



**Fig. 7.10** Spectrogram of a thunderstorm recorded in the Netherlands, depicting high-frequency (i.e., >8 kHz) sound from raindrops falling nearby, constant high-frequency (i.e., 9–12 kHz) rain in the background, and low-frequency (i.e., <1 kHz) sound from thunder

(*Strix aluco*) to 1/69th of the space without rain, with a simultaneous marked decrease in vocal activity. Thunder is the most common loud natural sound with a peak frequency near 100 Hz, although sounds extend into the infrasonic and mid-frequency range (250 Hz–4 kHz; Fig. 7.10). Other sources of terrestrial geophony are rivers, waterfalls, earthquakes, and volcanic eruptions. Infrasonic monitoring of soundscapes can identify the location of continuous geophonic sound sources, such as waterfalls and seismic activity, as well as transient (i.e., short-duration) sound sources, such as thunder, up to distances of 10 km (Johnson et al. 2006).

### 7.2.3 Anthropophony

Anthropophony identifies the presence and activities of human beings. Some of these sounds give cues about local culture, tradition, language, working habits, and religion (e.g., voices, music, cow and sheep bells, church bells, etc.) and can enrich a soundscape (Stack et al. 2011, Pavan 2017). However, with the industrial revolution, new sound sources have emerged at an unprecedented level and spatial extension, with consequent impacts on natural soundscapes and human health.

Terrestrial anthropophony includes sounds from transportation (e.g., road vehicles, trains, snowmobiles, ships, and airplanes; Ernstes and Quinn 2016; Mullet et al. 2017b; White et al. 2017; Duarte et al. 2019), recreational boats (Kariel 1990; Bernardini et al. 2019), machinery (e.g., excavation devices, drilling devices, generators, and chain saws; Potočnik and Poje 2010; Deichmann et al. 2017), gunshots (Wrege et al. 2017), fireworks (Kukulski et al. 2018), and outdoor events (Greta et al. 2019; Kaiser and Rohde 2013). The intensity of anthropophony correlates with the degree of urbanization (Joo et al. 2011; Kuehne et al. 2013) and is considered noise pollution with an impact on both human (European Environment Agency [EEA] 2014) and animal health (Barber et al. 2010; Shannon et al. 2016), potentially affecting entire ecosystems (Pavan 2017).

Low-frequency sound, mostly generated by engines, propagates over large distances and appears to be the most invasive and pervasive sound related to transportation infrastructures. Sound from cars and heavy trucks caused by tire-pavement interaction, aerodynamic sources, and engines peaks around 100 Hz (Rochat and Reiter 2016), but may reach as high as 10 kHz when measured close to the source (Fig. 7.11a). Both birds (e.g., Halfwerk and Slabbekoorn 2009) and anurans (e.g., Cunnington and Fahrig 2010; Caorsi et al. 2017) have been found to change vocal behavior in response to traffic noise (see Chap. 13). Conventional railway sound (i.e., electrified railway with a service speed $<$200 km/h) has a broad peak between 10 Hz and 2 kHz, whereas high-speed railway sound (i.e., electrified railway with a service speed $>$200 km/h) peaks $<$100 Hz (Di et al. 2014).

Sound from aircrafts, especially near airports, is perceived by humans as a source of disturbance and may have negative effects on children's learning, human sleep, and human health (Basner et al. 2017). In addition, sound during take-off and landing overlaps with biophony resulting in acoustic and behavioral responses (Fig. 7.11b; Sáncez-Pérez et al. 2013; Vidović et al. 2017). Birds near international airports in Spain, for example, were found to advance their dawn chorus to reduce overlap with aircraft sound (Gil et al. 2015), which is a common response to noise for urban species (Bermúdez-Cuamatzin et al. 2020). However, common chiffchaffs (*Phylloscopus collybita*) near airports in the UK and the Netherlands were found to sing songs with a lower maximum and peak frequency than conspecifics in nearby control areas, thus resulting in an increased overlap with aircraft sound (Wolfenden et al. 2019). In addition, airport populations sang at a slower rate and responded more aggressively to song playbacks. In South Africa, the critically endangered Pickersgill's reed frog (*Hyperolius pickersgilli*) called more frequently and at higher frequencies during and after aircraft overflights than before (Kruger and Du Preez 2016). Even in wild remote areas, aircrafts flying at ~8000 m altitude may

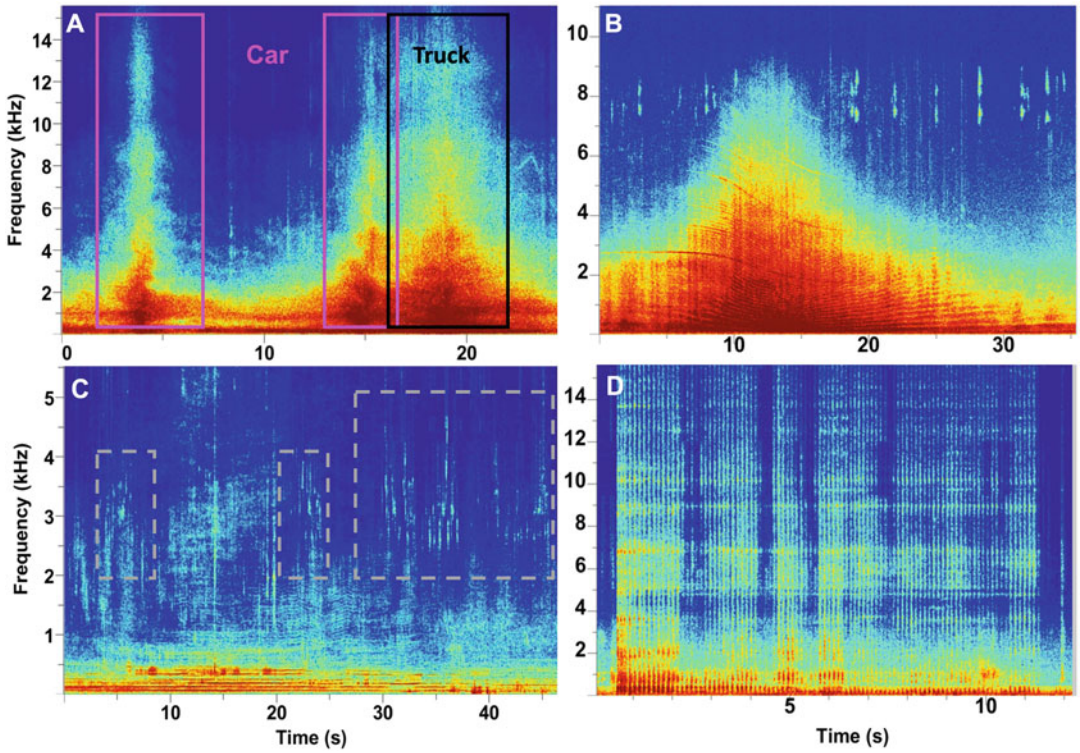



**Fig. 7.11** (**a**) Spectrogram of a passing car at 2-m and a truck at 5-m distance. (**b**) Spectrogram of a commercial passenger airplane flying overhead at an altitude of ~300 m after take-off. Note the Doppler shift from high to low frequency (from 2.8 to 2 kHz) around the time of closest approach (at ~12 s) and the bird vocalizations between 7 and 9 kHz. (**c**) Spectrogram of a 3-m recreational power boat with a 3-hp 2-stroke engine, passing at 5-m distance; bird vocalizations within the gray dashed boxes. (**d**) Spectrogram of a jackhammer breaking tar

produce noise below 500 Hz at 60 dB re 20 μPa (unweighted) at ground level (Pavan 2017; Farina et al. 2021). It is also essential to consider that take-off and landing corridors, where the noise levels are much higher, may cross more rural lands where airplane sound creates a stark contrast with ambient sound levels.

Smaller transport vehicles, such as powered two wheelers and snowmobiles, also contribute to the soundscape (Paviotti and Vogiatzis 2012; Mullet et al. 2017b). Mullet et al. (2017b) found that snowmobile noise, with main energy <2 kHz, affected 39% of the Alaskan wilderness open to snowmobiles and may mask vocalizations from common winter bird species. In-air ship noise from machinery and ventilation systems may propagate to areas near channels, ports, and coasts (Badino et al. 2012; Borelli et al. 2016). Small recreational power boats on lakes, on rivers, and near shore also increase in-air sound levels, predominantly below 1 kHz (Fig. 7.11c), with potential negative effects on bird species and hauled-out sea lions (York 1994; Tripovich et al. 2012).

Construction equipment may generate strong sounds that are audible over long ranges. Pneumatic tools, for example, generate repetitive, broadband sound (Fig. 7.11d). Heavy and stationary equipment, such as earth-moving machinery and air-compressors, generate sounds at frequencies <2 kHz (e.g., Berglund et al. 1996; Roberts 2009). Although one may associate construction sounds with urban areas, there are many examples in rural and remote areas, too. In the western Amazon (Peru), sounds from the construction and operation of a natural gas-well and

pipeline (i.e., generators, helicopters, and pneumatic tools) were audible up to 250 m from the source (Deichmann et al. 2017). Anthropogenic sources in rural areas include farming machinery dominating <500 Hz (Gulyas et al. 2002), chainsaws recorded in forests with main energy between 100 Hz and 9 kHz (Potočnik and Poje 2010), and transient, broadband gunshots (Prince et al. 2019), which can provide valuable information on illegal hunting, in particular in remote areas that are difficult to patrol. In urban settings, additional sources of anthropophony originate from outdoor events, such as (music) festivals (Greta et al. 2019), fun parks (Kaiser and Rohde 2013), and Formula 1 races (Payne et al. 2012).

### 7.2.4 Sound Propagation in Terrestrial Environments

The propagation of sound, from its source through an environment, affects the local soundscape. In environments with good sound propagation conditions, sources from far away contribute to the local soundscape; whereas in environments with poor sound propagation conditions, only nearby sources contribute. Sound propagation is affected by air temperature, humidity, ground cover (bare rock versus grasslands or bush), wind, turbulence, and the presence of sound absorbers (e.g., snow), scatterers (e.g., trees), and reflectors (e.g., cliffs or buildings; see Chap. 5).

As sound spreads, it is transmitted into and through different media, absorbed, reflected, scattered, and diffracted. Many of these effects depend on frequency; meaning that sound propagates differently at different frequencies and that the environment changes the spectral characteristics of the sound. If the wavelength of sound is smaller in size than features of the environment (e.g., rocks), then sound will reflect. The wavelength can be computed as the ratio of sound speed (about 330 m/s in air) and frequency (e.g., a 100-Hz tone has a wavelength of 3 m in air; see Chap. 4). At wavelengths much greater than features in the environment, sound will travel unhindered.

The air may be layered, with layers at different altitudes having different acoustic properties. Higher temperature and higher humidity increase the speed of sound. By Snell's law of refraction, sound bends toward the horizontal when the speed of sound increases and away from the horizontal when the speed of sound decreases. During the day, temperature typically decreases with increasing altitude, leading to an upward refracting environment that exhibits so-called shadow zones that have reduced sound levels. In the morning or in winter, the air near the ground is often relatively cold, while there might be a warmer layer of air at higher altitude; this situation is called a temperature inversion. Sound is downward refracted and channeled close to the ground. Hence, in winter, sound might travel very far at low altitude (see Chap. 5).

Vegetation attenuates sound, so in temperate areas with high vegetation, the same sound during summer propagates over shorter distances than during winter (Aylor 1972). Areas or seasons of full vegetative cover have soundscapes different from those bare in vegetation (Attenborough et al. 2012). Both temperature and humidity near the ground may change quickly; therefore, sound propagation conditions, soundscapes, and the communication space of terrestrial animals can vary within a few hours.

## 7.3 Aquatic Soundscapes

The vast majority of aquatic soundscape studies have focused on marine and estuarine environments, where soundscapes vary among geographic regions from the northern marginal ice-zone via equatorial regions to Antarctic waters (Haver et al. 2017), from the deep ocean (e.g., Dziak et al. 2017) to shallow coastal waters (e.g., McWilliam and Hawkins 2013), and from urban rivers (e.g., Marley et al. 2016) to estuarine reserves (e.g., Ricci et al. 2016). Soundscape studies in freshwater are less common but have covered a variety of settings from frozen lakes in Canada (Martin and Cott 2016) to urbanized lakes in the UK (Bolgan et al. 2016, 2018b), from pristine swamps in Costa Rica (Gottesman et al.

2020) to urbanized lowlands in the Netherlands (van der Lee et al. 2020), and from litttle streams in the USA (Holt and Johnston 2015) to the busy Ganges river in India (Dey et al. 2019). As in the terrestrial environment, each soundscape is characterized by a unique composition of biophony, geophony, and anthropophony.

Ambient sound encompasses all of the sounds at a given location and time, except for any specific signal of interest (International Organization for Standardization [ISO] 2017). Fig. 7.12 gives the spectra of characteristic ambient sounds in the ocean, as originally compiled by Wenz (1962), with updates from Cato (2008). Below 100 Hz, ambient sound is dominated by distant shipping, and, in shallow water, wind. Above 100 Hz, ambient sound is mostly wind driven. The prevailing limits of ambient sound decrease with increasing frequency from a maximum of 140 dB re 1 $\mu Pa^2$/Hz at 1 Hz to a minimum of 15 dB re 1 $\mu Pa^2$/Hz at 30 kHz. Above 30 kHz, molecular agitation limits the spectra of recorded ambient sound.

### 7.3.1 Biophony

Aquatic species are well adapted to produce, sense, and use sounds in water (e.g., Schmitz 2002; Ladich and Winkler 2017). The aquatic biophony includes sounds produced by invertebrates (e.g., Iversen et al. 1963; Coquereau et al. 2016; Gottesman et al. 2020), frogs (Brunetti et al. 2017), turtles (e.g., Giles et al. 2009), fish (e.g., Kasumyan 2008; Bolgan et al. 2018b), birds (Thiebault et al. 2019), and mammals (e.g., Klinck et al. 2012; Erbe et al. 2017; Dey et al. 2019). The freshwater biophony is not well described and so, sounds frequently cannot be linked to specific species (Rountree et al. 2019; Gottesman et al. 2020; Putland and Mensinger 2020). This lack of knowledge currently impedes the full utilization of freshwater soundscape studies as an ecological tool (Linke et al. 2020).

With regards to marine biophony, snapping shrimps are well-known contributors, producing broadband sounds from a few hundreds of hertz

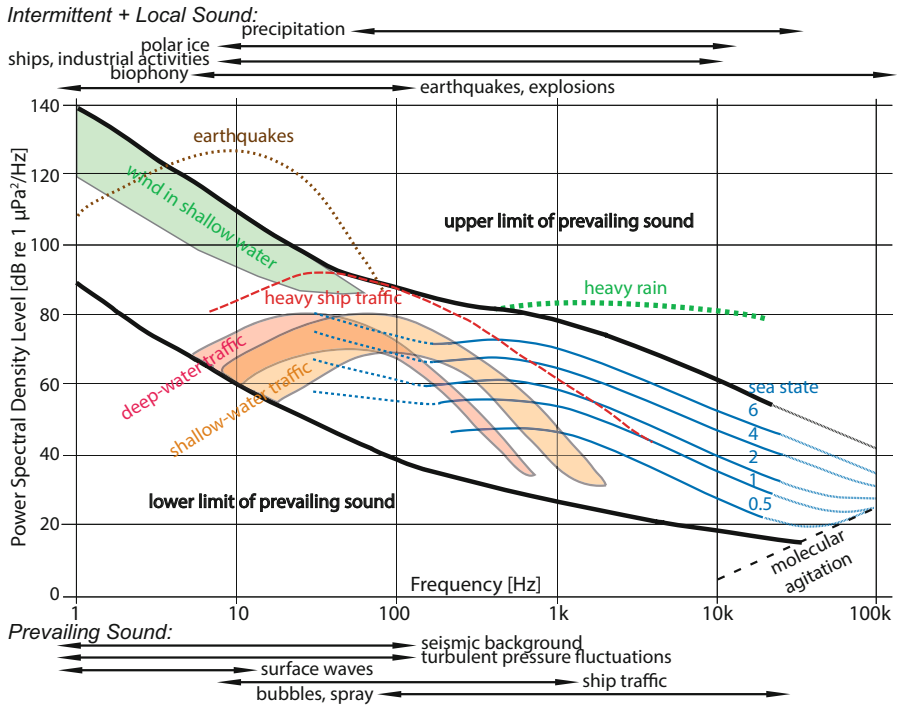



**Fig. 7.12** Spectra of prevailing and local underwater sound sources between 1 Hz and 100 kHz (after Wenz 1962; Cato 2008)

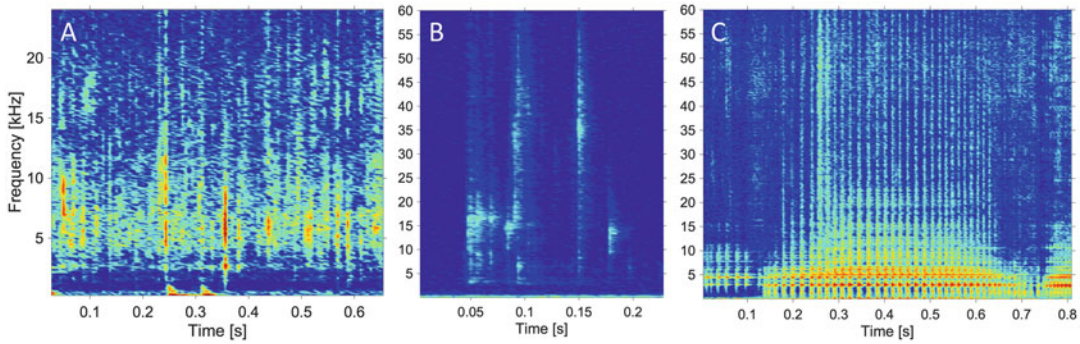



**Fig. 7.13** Spectrograms of (**a**) snapping shrimp, (**b**) a swimming great scallop (*Pecten maximus*), and (**c**) a feeding spider crab (*Maja brachydactyla*). Spectrograms **b** and **c** were created from supplementary material in Coquereau et al. (2016). Reprinted by permission from Springer Nature. Coquereau L, Grall J, Chauvaud L, et al. Sound production and associated behaviours of benthic invertebrates from a coastal habitat in the north-east Atlantic. Mar Biol 163: 127; https://doi.org/10.1007/200227-016-2902-2. © Springer Nature, 2020. All rights reserved

up to 200 kHz (Fig. 7.13a; Knowlton and Moulton 1963; Au and Banks 1998). This short, intense, repetitive sound is a byproduct of many shrimps rapidly closing their snapper claw, which creates a jet stream used in agonistic encounters and to stun prey (Herberholz and Schmitz 1999). As snapping shrimps predominantly live in large aggregations (Duffy 1996; Duffy and Macdonald 1999), their sounds can be heard as a constant 'crackling' chorus with temporal and spatial variations in intensity (e.g., Bohnenstiehl et al. 2016; Lillis et al. 2017). Other well-known sound-producing invertebrates are lobsters and sea urchins. Lobsters produce broadband pulse trains when facing predators or competing conspecifics (Staaterman et al. 2010; Jézéquel et al. 2019). Jézéquel et al. (2019) characterized pulse trains of the European spiny lobster (*Palinurus elephas*) as signals with a mean bandwidth of 5–23 kHz. Sea urchins scrape algae from rocks. This foraging strategy causes the fluid inside the sea urchin to resonate, producing sound at frequencies between 700 Hz and 2 kHz (Radford et al. 2008). In New Zealand, groups of foraging endemic Kina sea urchins (*Evechinus chloroticus*) increase sound levels between 18:00 and 20:00 compared to mid-day levels (Radford et al. 2008). Further examples of sounds from invertebrate movement and foraging activities are displayed in Fig. 7.13b, c (Coquereau et al. 2016).

Over 1200 fish species were estimated to produce sounds by Kaatz (2011), of which 800 were confirmed soniferous species (Kaatz 2002; Rountree et al. 2006). Fish produce sounds in a variety of behavioral contexts, such as courtship (Amorim et al. 2015), agonistic interactions (Ladich 1997), and when in distress (Knight and Ladich 2014). It is therefore not surprising that fish are common contributors to aquatic soundscapes, most noticeably when large numbers vocalize in chorus (e.g., Rice et al. 2017; Pagniello et al. 2019). Parsons et al. (2016) summarized fish chorus patterns over a 2-year period in Darwin Harbour, Australia. Nine different chorus types were detected (Fig. 7.14), dominating the frequency band from 50 Hz to 3 kHz and displaying cycles on several temporal scales (i.e., diurnal, lunar, seasonal, and annual). Fish chorusing was also associated with environmental parameters, including water temperature, depth, salinity, and tidal cycle.

Marine mammal sounds range from infrasounds of mysticetes (baleen whales; e.g., Mellinger and Clark 2003) to ultrasounds of odontocetes (toothed whales; e.g., Hiley et al. 2017). Calls may function as contact or warning signals. For example, northern right (*Eubalaena*

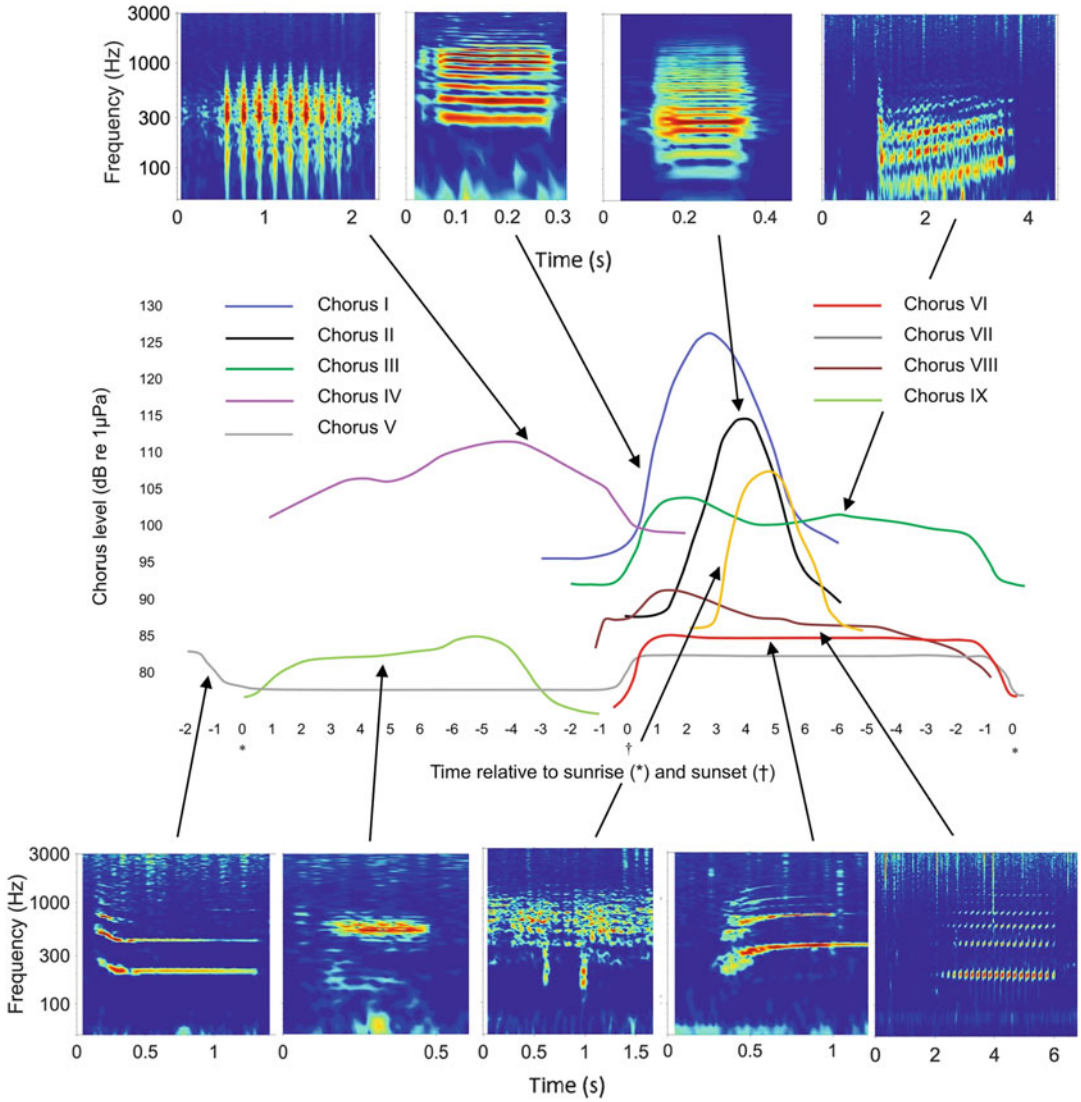



**Fig. 7.14** Spectrograms of the fish calls making up nine fish choruses (50 Hz–3 kHz) in Darwin Harbour, Australia. The middle panel shows the chorus levels over time, in hours relative to sunrise and sunset. There is a peak in chorusing activity shortly after sunset. Figure created from material in Parsons et al. (2016), by permission from Oxford University Press. Parsons MJG, Salgado-Kent CP, Marley SA, et al., Characterizing diversity and variation in fish choruses in Darwin Harbour. ICES J Mar Sci 73:2058–2074; https://doi.org/10.1093/icesjms/fsw037. © International Council for the Exploration of the Sea, 2016; https://global.oup.com/academic/rights/permissions/. All rights reserved. Reuse requires permission from OUP

*glacialis*) and southern right (*E. australis*) whale upsweeps (i.e., upcalls; 50–235 Hz) seem to be used as a contact call (Fig. 7.15a; Clark 1982; Parks et al. 2007). Another characteristic call of this species is a strong, brief, broadband pulse with energy up to 16 kHz (called gunshot), which may serve as an advertisement call and/or agonistic call produced by male individuals (Parks et al. 2006). However, female right whales sometimes also produce this sound (Gerstein et al. 2014). Foraging humpback whales (*Megaptera novaeangliae*) produce a characteristic tonal call

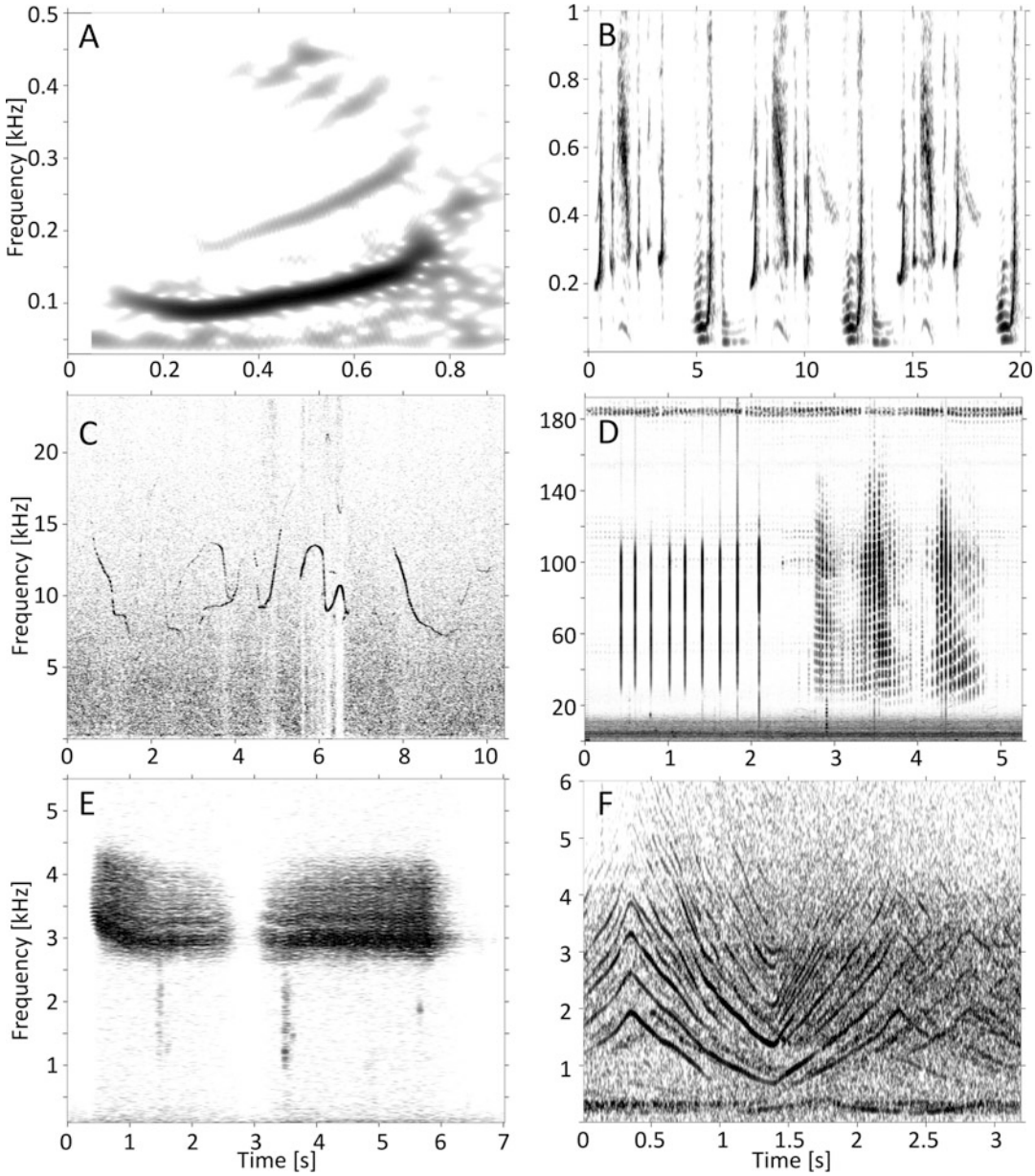



**Fig. 7.15** Spectrograms of marine mammal sounds. (**a**) Southern right whale upcall. (**b**) Humpback whale song. (**c**) Common dolphin (*Delphinus delphis*) whistles and (**d**) clicks and burst-pulse sounds. (**e**) Leopard seal (*Hydrurga leptonyx*) and (**f**) Ross seal (*Ommatophoca rossii*), both under water. © Erbe et al. (2017); https://doi.org/10.1007/s40857-017-0101-z. Published under CC BY 4.0; https://creativecommons.org/licenses/by/4.0/

with a fundamental frequency between 400 Hz and 1 kHz (Cerchio and Dahlheim 2001), which may function to herd prey, coordinate group movement, or recruit individuals into a feeding group (Cerchio and Dahlheim 2001; Fournet et al. 2018).

Blue whales (*Balaenoptera musculus*), bowhead whales (*Balaena mysticetus*), fin whales (*Balaenoptera physalus*), and others arrange calls into patterned song, which may last from hours to days. Humpback whale song is particularly complex in structure, consisting of a variety of units

that have peak frequencies between 20 Hz and 6 kHz (Fig. 7.15b; Payne and McVay 1971). The functions of whale song may include female attraction, male-male interactions, and long-range sonar (Herman 2017; Mercado 2018). Odontocete echolocation clicks with peak energy between ~10 and ~150 kHz are used for navigation and prey capture (Au 1993). Odontocete tonal calls (i.e., whistles) with fundamental frequencies between ~1 and ~50 kHz and broadband burst-pulse sounds are used for communication (Fig. 7.15c, d; Herzing 1996). Some odontocete species also communicate with clicks (e.g., sperm whales, *Physeter macrocephalus*, and porpoises, Phocoenidae; Weilgart and Whitehead 1993; Clausen et al. 2010). Delphinids may arrange their whistles and burst-pulse sounds into patterned sequences (e.g., killer whales, *Orcinus orca*, Wellard et al. 2020; and pilot whales, *Globicephala melas*, Courts et al. 2020). Seals, sea lions, and walruses use underwater vocalizations particularly during the breeding season and in social interactions (Schusterman et al. 1966; Stirling et al. 1987; Van Parijs and Kovacs 2002). The majority of pinniped underwater vocalizations fall within the frequency range between 10 Hz and 6 kHz (Fig. 7.15e, f), although Weddell seals (*Leptonychotes weddellii*) were found to produce calls containing energy up to 13 kHz (Thomas and Kuechle 1982). Mysticetes, odontocetes, and pinnipeds also produce non-vocal surface-generated sounds through breaching, pectoral fin slapping, and tail slapping (e.g., Dunlop et al. 2007).

### 7.3.2 Geophony

The aquatic geophony comprises sounds from wind acting on the water surface (e.g., Knudsen et al. 1948); precipitation (e.g., Nystuen 1986); ice movement, pressure cracking, and melting (e.g., Mikhalevsky 2001; Martin and Cott 2016); subsea volcanoes and earthquakes (e.g., Fox et al. 2001; Dziak and Fox 2002); and sediment displacement (e.g., Lorang and Tonolla 2014). Geophony can be nearly continuous and dominate the soundscape in certain regions at certain times (e.g., wind noise in southern Australia; Erbe et al. 2021). Wind-driven sound lies between 100 Hz and 20 kHz (typical peak at 500 Hz; Wenz 1962). Rainfall can contribute to the underwater soundscape over frequencies between 500 Hz and 50 kHz depending on drop size, rainfall rate, and impact angle related to wind speed (Ma et al. 2005). In the Perth Canyon, Australia, rainfall is often accompanied by strong wind. Consequently, the weather-related sound spectrum shows two peaks: one dominated by wind at 300–600 Hz and another dominated by rain at about 3 kHz (Fig. 7.16a; Erbe et al. 2015). In polar regions and underneath frozen lakes,

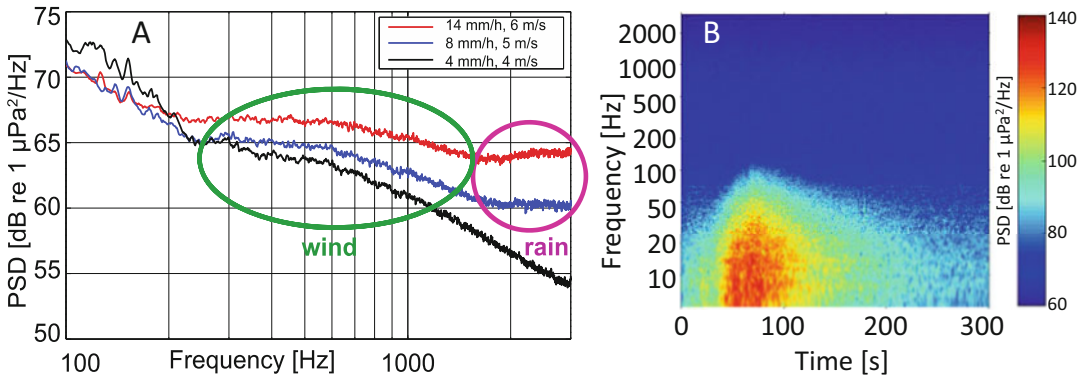



**Fig. 7.16** Sources of aquatic geophony. (**a**) Underwater power spectral density (PSD) levels illustrating an increase in levels under increased wind speeds (m/s) and rain fall rates (mm/h). (**b**) Spectrogram of an earthquake recorded in the Perth Canyon, Australia. Colors indicate PSD level (dB re 1 µPa$^2$/Hz). Note the logarithmic frequency axes. Both figures were modified; © Erbe et al. (2015); https://doi.org/10.1016/j.pocean.2015.05.015. Published under CC BY 4.0; https://creativecommons.org/licenses/by/4.0/

sounds of colliding, oscillating, breaking, and melting ice range from <10 Hz to 8 kHz (Talandier et al. 2006; Martin and Cott 2016). Sound from polar ice can be detected thousands of kilometers away at tropical latitudes (Matsumoto et al. 2014). Underwater volcanic eruptions generate impulsive sounds as well as harmonic tremors <100 Hz, which can travel over distances greater than 12,000 km through the Sound Fixing And Ranging (SOFAR) channel (Tepp et al. 2019). Similarly, earthquakes can be detected at thousands of kilometers in distance as low-frequency (<100 Hz) rumbles, lasting several minutes (Fig. 7.16b; Erbe et al. 2015). Sediment flow may generate sound in rivers and streams, creating acoustic cues for freshwater species (Tonolla et al. 2010, 2011). Depending on grain size and flow velocity, the spectrum may range from tens of hertz to kilohertz.

### 7.3.3 Anthropophony

In the last century, human activities began to contribute significantly to underwater sound levels. The anthropophony has grown ambient sound levels rapidly compared to evolutionary time scales, making it hard for animals to adapt (see Chap. 13). Anthropogenic sound may be present in aquatic soundscapes far away from human activities, owing to the long-range propagation of low-frequency sound in water (see Chap. 6). The aquatic anthropophony includes personal watercrafts (e.g., jetskis; Erbe 2013), small boats (e.g., Erbe et al. 2016a; Dey et al. 2019), electric ferries (Parsons et al. 2020), merchant ships (e.g., Ross 1976; Hatch et al. 2008; McKenna et al. 2012), offshore hydrocarbon exploration and production (e.g., marine seismic surveys and drilling; Wyatt 2008; Erbe and King 2009; Erbe et al. 2013), near-shore construction including geotechnical work and pile-driving (e.g., Erbe 2009; Dahl et al. 2015; Erbe and McPherson 2017), windfarms (e.g., Koschinski et al. 2003; Tougaard et al. 2009), dredging (e.g., Reine et al. 2014), explosions (e.g., Soloway and Dahl 2014), military sonars (e.g., Ainslie 2010), acoustic alarms on fishing gear or shark nets (e.g., Erbe and McPherson 2012), snowmobiles and vehicles on ice-covered lakes (Martin and Cott 2016), bridge traffic (Holt and Johnston 2015; Martin and Popper 2016), augers (i.e., ice drills; Putland and Mensinger 2020), airplanes (e.g., Martin and Cott 2016; Erbe et al. 2018), and activities alongside, rather than on, the water (Kuehne et al. 2013). Lesser-known anthropophony originates from unpowered recreational activities (e.g., scuba diving and swimming; Erbe et al. 2016c).

Sound from ship traffic is the most pervasive anthropogenic sound in the ocean (e.g., Sertlek et al. 2019). The level of sound emitted depends on ship type, size, speed, and operational mode (e.g., reversing, idling, carrying, or towing load; MacGillivray and de Jong 2021). In water <300 m deep, large ships (>300 t) can temporarily increase sound levels up to 125 kHz within 500 m from shipping routes (Hermannsen et al. 2014; Veirs et al. 2016). In deep water, low-frequency sound from ships can travel farther, especially when entering the SOFAR channel (Fig. 7.17; Erbe et al. 2019). The number of small, recreational boats that occupy coastal waters is on the rise in many places and these vessels may raise sound levels between 100 Hz and 20 kHz in coastal and estuarine habitats, depending on boat type, hull type, length, propulsion system, operational mode, and speed (Parsons et al. 2021).

Another common anthropogenic sound that has received much concern over its potential impacts on marine life (see Chap. 13) is produced by seismic surveys, used for seabed profiling and hydrocarbon exploration. Surveys are done with a vessel towing an array of airguns. Airguns are metal chambers storing compressed air, which is rapidly released, producing an acoustic pulse with energy up to at least 10 kHz (Dragoset 2000; Hermannsen et al. 2015). Airguns exist with different operating volumes and firing pressures, affecting the spectrum and level of the acoustic pulses (Fig. 7.18a; Erbe and King 2009; Hermannsen et al. 2015). Airgun arrays can be tuned to focus acoustic emission down into the seabed, yet some sound ends up traveling horizontally through the water. Hence, sounds from

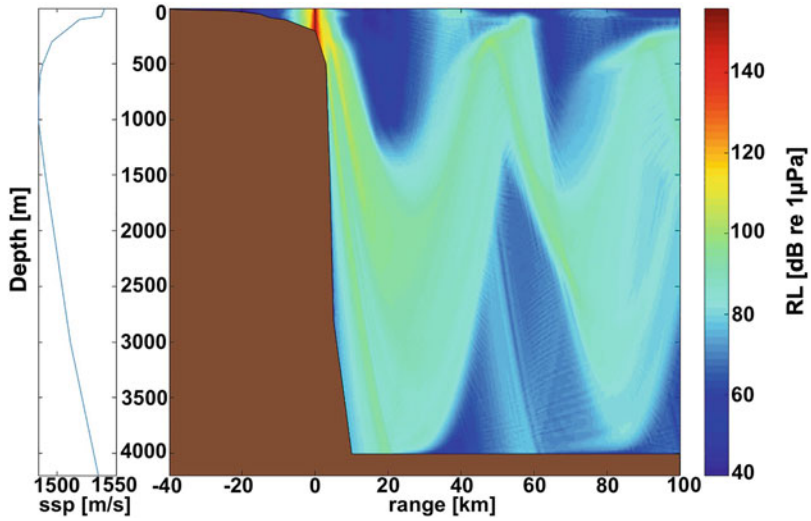



**Fig. 7.17** Sketch of the propagation of sound from a 156-m ship (at 0 km range) sailing at a speed of 15 knots above the continental slope in the absence of ambient sound. Propagation modeled with *RAMGeo* in *AcTUP V2.8* (https://cmst.curtin.edu.au/products/underwater/) with an equatorial sound speed profile as indicated in the left panel and a hard, dense, limestone seafloor. Colors represent received level (RL). © Erbe et al. 2019; https://www.frontiersin.org/files/Articles/476898/fmars-06-00606-HTML/image_m/fmars-06-00606-g001.jpg. Published under CC BY 4.0; https://creativecommons.org/licenses/by/4.0/

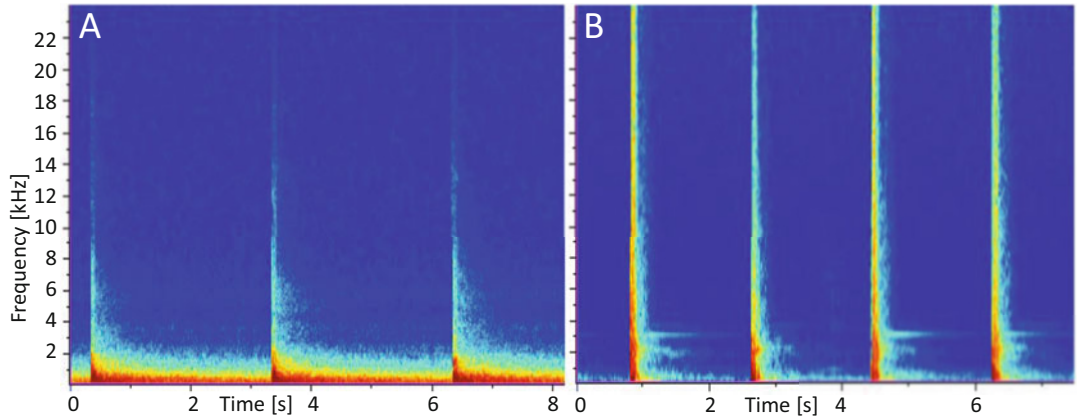



**Fig. 7.18** Spectrograms of impulsive sound sources. (**a**) Seismic airgun pulses recorded off Western Australia (Erbe et al. 2021). (**b**) Pile driving recorded in Moreton Bay, Queensland, Australia (Erbe 2009)

seismic surveys may affect marine life at both short and long ranges (Gordon et al. 2003; Slabbekoorn et al. 2019). A typical seismic survey may last several weeks, during which the airgun array is discharged every few seconds.

Other common sounds of concern are emitted by pile driving, explosions, and acoustic alarms. Pile driving for windfarm construction and detonations of World War II ammunition are regular sources of sound within European waters (Bailey et al. 2010; von Benda-Beckmann et al. 2015). Impact pile driving generates high-intensity pulses with energy exceeding 40 kHz at close range (Fig. 7.18b). Acoustic alarms are

devices that purposefully emit sound between a few hundred hertz and tens of kilohertz to deter marine animals from potential hazards, such as pile driving sites, aquaculture farms, or bather protection nets (e.g., Jacobs and Terhune 2002; Erbe and McPherson 2012), yet their efficacy remains controversial (e.g., see Erbe et al. 2016d). Acoustic alarms differ widely in their signal type, frequency, and source level (Findlay et al. 2018).

### 7.3.4 Sound Propagation in Aquatic Environments

Underwater, the propagation of sound is affected by water temperature, salinity, hydrostatic pressure (i.e., depth below the sea surface), sea surface roughness, potential ice cover, bathymetry, seafloor roughness, upper seafloor geology (i.e., sediment type and thickness), depth and type of the underlying bedrock, and the presence of sound absorbers, scatterers, and reflectors (e.g., aquatic fauna, bubble clouds, or suspended sediment; see Chap. 6).

The speed of sound in water changes gradually with depth. As a result, sound does not travel in straight lines. Instead, sound paths are bent by refraction. By Snell's law, paths bend toward local minima in sound speed. The most pronounced local minimum occurs in all non-polar oceans at a depth of about 1000 m below the sea surface. Sound reaching this depth at not too steep angles can get trapped in the so-called SOFAR channel by being repeatedly refracted toward the channel axis. This is how sound can traverse entire oceans, with sound sources contributing to soundscapes thousands of kilometers away (e.g., Gavrilov 2018). The SOFAR channel does not only trap sounds from deep-water sound sources (e.g., submarines or diving megafauna) located within the channel, but also from sources near the sea surface (e.g., ships or whales) because sound can radiate into the SOFAR channel with just one reflection off a downward sloping seafloor (Fig. 7.17). The minimum in sound speed (and so the axis of the SOFAR channel) rises to shallower depths in polar waters. In fact, in the polar oceans, the speed of sound is the smallest at the surface. This leads to a surface duct, in which sound travels by repeated reflection off the sea surface and refraction at depth.

Snell's law creates additional interesting phenomena such as shadow zones and convergence zones. Sound does not distribute evenly throughout the oceans. There are patterns of shadow zones (into which sound cannot travel by direct paths, and which receive little to no sound) and convergence zones (where received levels are enhanced; Fig. 7.17). These zones will be in different places for different source locations. In addition, sound at low frequencies does not travel far in shallow water. The waveguide concept and normal modes nicely explain this (see Chap. 6). The water depth can be too small to "fit" sound of large wavelength. As a result, ship noise may be attenuated quickly in coastal water and the spectral hump of distant shipping is characteristic only in offshore water (see Sect. 7.5.3.2). Ergo, soundscapes may differ with location and depth, merely because of sound propagation.

## 7.4 Soundscape Changes Over Space and Time

Soundscapes may vary on a range of spatial scales, exhibit temporal cycles (e.g., because of diurnal animal behaviors, periodic animal presence, or seasonal weather events; Erbe et al. 2015; Caruso et al. 2017; McWilliam et al. 2017), or gradually change over longer periods of time. Such changes may be natural or, directly or indirectly, related to human activity. Understanding natural variability is important for using soundscapes (1) as an ecological tool to study animal behavior and (2) as a management tool of the potential effects of human activity. Our understanding of the function of animal calls and natural or anthropogenic interferences is based on limited observational data (Slabbekoorn et al. 2018) and so interpreting changes in sounds is even more difficult. Gavrilov et al. (2012), for example, recorded the underwater soundscape between 21 and 27 May in 2002, 2006, and 2010 off Cape Leeuwin, Australia. Between

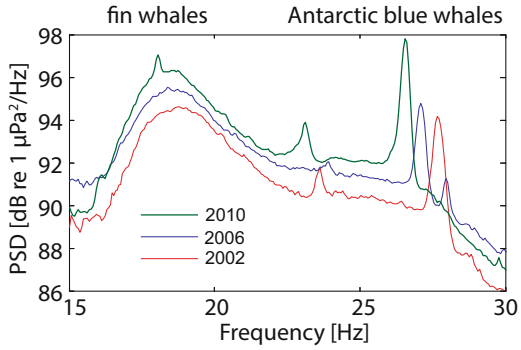



**Fig. 7.19** Power spectral density (PSD) of the soundscape off Cape Leeuwin, Australia, showing increases in level and decreases in frequency of the fin and Antarctic blue whale characteristic sounds over eight years. Figure courtesy of Sasha Gavrilov, Curtin University, Perth, Australia

years, an increase in sound levels at the frequencies characteristic of fin whales and Antarctic blue whales (*Balaenoptera musculus intermedia*) was seen (Fig. 7.19). This could be due to an increase in whale population sizes or changes in migration routes (i.e., closer to the recorder). The authors further noted that the frequency of Antarctic blue whale calls decreased for unknown reasons.

### 7.4.1 Spatial Patterns

Soundscapes vary naturally over large and small spatial scales, abruptly or gradually, resulting in different soundscapes between and within habitats. Slabbekoorn (2004) sampled multiple sites within a contiguous rainforest and an adjacent ecotone forest in Cameroon. He found spatial differences in ambient noise, which were due to differences in wind and species vocalizations (insects, frogs, and birds). Over time, ambient noise can affect the vocal characteristics of individuals, populations, and species (see Chap. 13). Consistent ambient noise may drive the features of a species' vocalizations, so that call transmission is optimized within the acoustic environment (Acoustic Adaptation Hypothesis). Just as temporal changes in ambient noise may result in vocalization changes, spatial changes in ambient noise may result in spatial differences in vocalizations (Slabbekoorn and Smith 2002). If ambient noise differs consistently across a species' habitat, acoustic adaptation might result in acoustic divergence between populations of the same species (Dingle et al. 2008). If the calls of these populations diverge so much that they are no longer recognized by all populations, sexual selection may lead to the segregation into distinct (sub)species (Dingle et al. 2010; Burbidge et al. 2015). For research on soundscapes and acoustic ecology, spatial replication in sampling is paramount.

### 7.4.2 Natural Cycles

Soundscapes vary naturally with diurnal, lunar, seasonal, or annual cycles because of temporal patterns in animal presence and behavior (e.g., night-time foraging, lunar spawning, seasonal hibernation, and annual migration) as well as weather (e.g., annual monsoon). In Alaska, ambient sound increased rapidly in early spring due to an influx of migratory bird species and the awakening of species from dormancy and hibernation (Mullet et al. 2016). Gage and Axel (2014) studied the diurnal and seasonal patterns in ambient sound within 1-kHz frequency bands at Michigan Lake, USA, from 2009 to 2012. At 2–3 kHz, power levels were highest in early spring with the presence of spring peepers (*Pseudacris crucifer*, Hylidae). Levels dropped progressively toward early fall when spring peepers disappeared and increased again in late fall because of chorusing insects. In contrast, at 4–5 kHz, levels were low in early spring but increased in late spring with the presence of breeding birds. Levels subsequently dropped yet increased again in late summer and early fall because of insects. Diurnal changes in ambient sound were related to ecological activity. Within the 2–4 kHz frequency band, for example, spring peepers dominated the soundscape at night until singing birds took over at dawn. Under water, in the Ionian Sea, echolocation activity of dolphins occurred at nighttime and crepuscular hours (Caruso et al. 2017). In contrast, communication

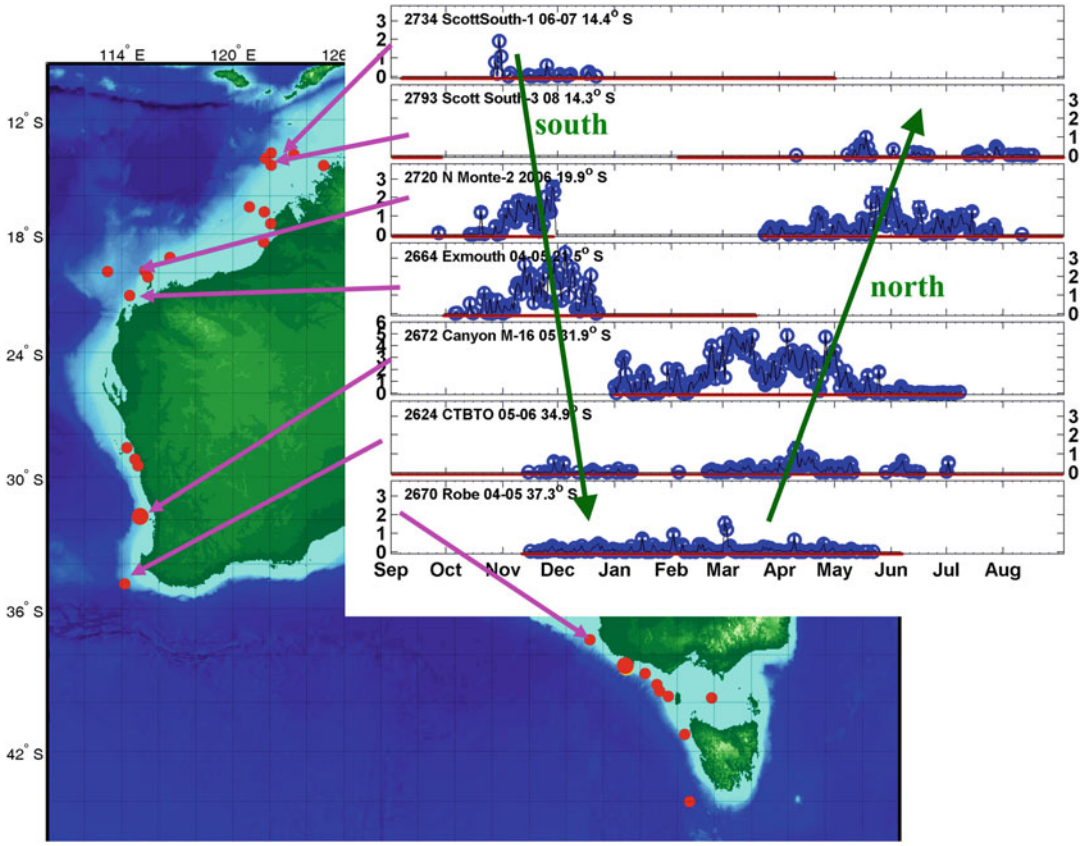



**Fig. 7.20** Seasonal timing of pygmy blue whale migration along the west and south coasts of Australia based on passive acoustic monitoring. The chart shows the locations of sound recordings (red dots). The diagram shows counts of pygmy blue whale singers as 24-h means. The red horizontal lines indicate when the recorders were operating (Erbe et al. 2016b)

signals (i.e., whistles) were mostly produced during the day. Seasonal variation, with a peak number of clicks in August, was also evident, but no effect of lunar cycle was observed. Off Western Australia, pygmy blue whales (*Balaenoptera musculus brevicauda*) are a seasonally dominant contributor to the marine soundscape and simply by listening, their seasonal migration can be traced along the coast (Fig. 7.20; Erbe et al. 2016b).

## 7.4.3 Human Activities

In many habitats, soundscapes have changed significantly over the last century, with habitat degradation by humans as a root cause. Humans add sound to soundscapes, change biodiversity through land-use, and directly remove animals from habitats (e.g., by hunting). Humans also contribute to climate change, with greenhouse gas emissions resulting in environmental changes, which can have direct and indirect effects on ecosystems and related soundscapes. The conservation of soundscapes is important not only for scientific and ecological reasons but also for touristic interests and human welfare (Pavan 2017).

### 7.4.3.1 Anthropophony

Humans alter soundscapes by growing anthropophony through an increase in transportation, construction, mineral and hydrocarbon

exploration and production, military exercises, recreational activities, etc. These activities produce sounds over a wide range of frequencies and at a variety of intensities (see Sects. 7.2.3 and 7.3.3). While some activities are temporary, others result in sustained increases in ambient sound levels over time. For example, underwater sound from shipping has increased ambient sound levels between 10 and 100 Hz in large parts of the world's oceans by up to 3 dB per decade (e.g., Andrew et al. 2011; Chapman and Price 2011; Miksis-Olds et al. 2013).

Seismic surveys produce intense sound over a few weeks at a time to explore a specified area; yet, Nieukirk et al. (2004, 2012) detected airgun pulses along the Mid-Atlantic ridge from seismic survey vessels located 3000–4000 km away. In 1999, airgun signals were routinely detected for more than 80% of the days in a month, which increased to 95% in 2005. Finally, anthropogenic sounds may affect animal behavior (i.e., physical or acoustic, Slabbekoorn et al. 2018; see Chap. 13), which can further alter soundscapes.

#### 7.4.3.2 Land Use

Humans transform natural landscapes to increase agricultural land coverage, to build infrastructure (e.g., roads, buildings, and power supply systems), or to extract resources (e.g., tree logging and mining). These activities generate sound and affect animal density and biodiversity, ultimately changing soundscapes (Phillips et al. 2017). In 1962, ecologist Rachel Carson expressed her concern about the use of chemicals and pesticides in agriculture, killing not only soil micro-fauna but also macro-fauna (Carson 1962). She foresaw a silent natural world without the songs of insects, frogs, and birds, if they were lost due to urbanization or chemical pollution. She was one of the first to consider animal sounds as an expression of ecosystem integrity and quality. Kerr and Cihlar (2004) found a correlation between high-intensity, high-biomass agriculture and high numbers of endangered species on both national and regional levels in Canada.

Danielsen and Heegaard (1995) compared the species richness and abundance of birds, primates, squirrels, tree-shrews, and bats between undisturbed, logged, and transformed patches of forest (i.e., to rubber and oil palm plantations) in eastern Sumatra, Indonesia. Logging changed the composition of bird species, revealing a decrease in the number of specialized insectivorous species and an increase in insectivore-frugivore generalist species. The species richness of bats also decreased with a concomitant increase in abundance of the most dominant bat species. However, logging impacts differed between geographical regions and management strategies (e.g., conventional selective, salvage, or reduced-impact logging; Chaudhary et al. 2016; LaManna and Martin 2017). Land transformation to plantations resulted in a dramatic decrease in biodiversity with the disappearance of primates, squirrels, and tree-shrews as well as a reduction in bird and bat species richness by 90–95% and 75–87%, respectively.

#### 7.4.3.3 Direct Takes

Accidental, illegal, or over-harvesting of animal species occurs in both terrestrial and aquatic habitats (e.g., Challender and MacMillan 2014; Anderson et al. 2020), resulting in population declines and species extinctions (Hoffmann et al. 2011; Dulvy et al. 2014). Perhaps one of the greatest examples is the removal of millions of whales during the nineteenth and twentieth centuries (Rocha Jr. et al. 2014), which unequivocally changed marine soundscapes world-wide. A modern example is the threat of dissapearing Gulf corvina (*Cynoscion othonopterus*) choruses in the Colorado River delta because of overfishing (Erisman and Rowell 2017). Overfishing can also result in excessive growth of algae, ultimately changing soundscapes. Freeman et al. (2018), for example, found a positive correlation between sound levels and macroalgae coverage on Hawaiian coral reefs, attributable to ringing bubbles emitted during photosynthesis.

#### 7.4.3.4 Climate Change

The Earth is experiencing rapid climate change, affecting soundscapes in a variety of ways. The geophony is affected by changing weather

patterns (i.e., wind, precipitation, and storms; Sueur et al. 2019). Rising temperatures reduce sea- and land ice, which is changing polar soundscapes (Intergovernmental Panel on Climate Change [IPCC] 2014). Climate change further modifies the acoustic properties of the environment with direct effects on sound propagation and thus the audible distances of sounds. Larom et al. (1997) calculated that the effective communication range for African elephant calls varied between 2 and 10 km with temperature and windspeed. Ocean acidification, as a result of climate change, results in less absorption of low-frequency sounds (Gazioğlu et al. 2015). Thus, low-frequency sound sources, such as ships and whales, may become more prominent in future marine soundscapes.

Climate change may also directly affect a species' vocal behavior, distribution pattern, or timing of behavioral events, such as migration and mating (Krause and Farina 2016; Sueur et al. 2019). Narins and Meenderink (2014) found that Puorto Rican coqui frogs (*Eleutherodactylus coqui*), over a period of 23 years, moved to higher altitudes, while their calls increased in pitch and decreased in duration. These changes in distribution and call characteristics corresponded to an overall increase in temperature of 0.37 °C, with a concomitant decrease in body size. A different response was seen by four frog species near Ithaca, NY, USA, who advanced the start of their breeding season by 14 days between 1900–1912 and 1990–1999, as evident from recordings of mating calls (Gibbs and Breisch 2001). During this time, temperatures increased on average 0.7–1.7 °C. Insects also depend on air temperature for the expression of their behavior, including sound emission (Ciceran et al. 1994). Rossi et al. (2016a, b) found that snapping shrimp (family Alpheidae) reduced their snap rate (i.e., snaps per minute) and intensity under increased levels of $CO_2$. This might affect the behavior of species that rely on acoustic cues from snapping shrimp for navigation (Rossi et al. 2016b). The eastern Chukchi Sea beluga whale (*Delphinapterus leucas*) population delayed timing of migration from foraging habitats by 2–4 weeks, corresponding to a delay in regional sea-ice freeze-up (Hauser et al. 2016). These examples stress the importance of collecting environmental data together with acoustic data, to correlate changes in animal distribution patterns and behavior with environmental change (Kloepper and Simmons 2014).

## 7.5 How to Analyze Soundscapes

Soundscape analysis may involve various, sometimes sequential, methods ranging from listening to recordings, via visual inspection of spectrograms, to automated detection of target signals, and computation of several acoustic metrics. Often, the larger the acoustic monitoring project, the more automated the tools, as long-term projects, which might compare multiple recording sites, might gather terabytes of data, which are virtually impossible to analyze by hand.

### 7.5.1 Standard Soundscape Measurements

Initial assessments of soundscapes typically involve the computation of spectrograms and some general statistics, such as the broadband root-mean-square (rms) Sound Pressure Level ($SPL_{rms}$) in either dB re 20 μPa or dB re 1 μPa in air and water, respectively (see Chap. 4). This allows an initial quality-check of the recordings and the identification of potential spatial or temporal patterns in overall sound levels, highlighting areas or temporal events of interest for further investigation (e.g., very quiet or very noisy areas or times of day, Fig. 7.21). However, broadband $SPL_{rms}$ levels are strongly influenced by the noisiest events and cannot identify the myriad of soundscape components and contributors to spatial and temporal differences.

As sound sources are often known to cover certain frequency bands, it is beneficial to compute SPLs within purposefully chosen frequency bands or standard octave or 1/3 octave bands. Buscaino et al. (2016) used Octave Band Levels (OBLs) at center frequencies from 62.5 Hz to 64 kHz to study temporal patterns in the

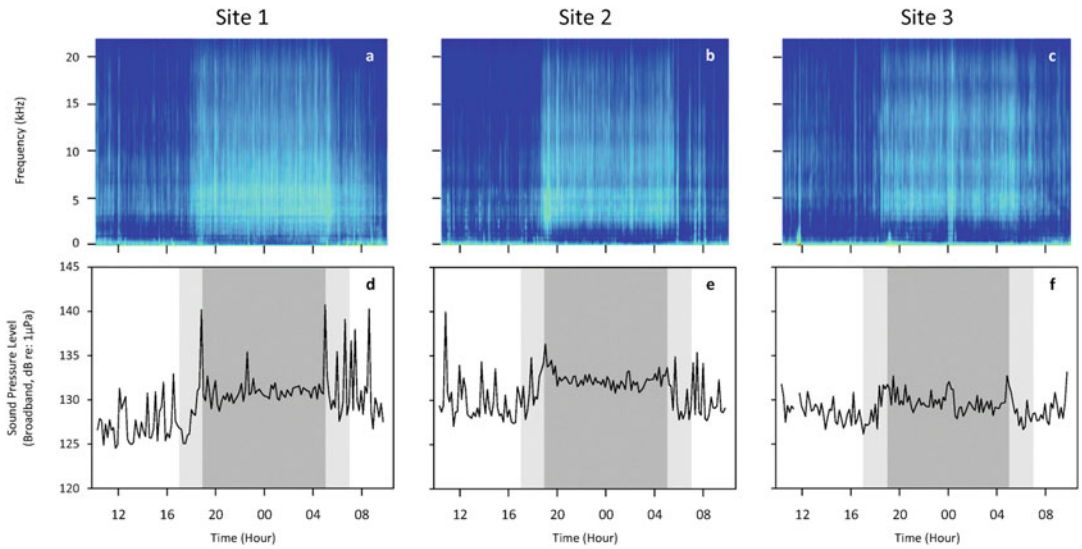



**Fig. 7.21** Spectrograms (top) and time series (bottom) of broadband (20 Hz–22 kHz) sound pressure levels of a 24-h recording period at three sites around Bora Bora Island, French Polynesia. Recording schedule was set at 60 s every 10 min. Note the increase in sound levels at night (shaded areas) as well as the strong fluctuation in sound levels between 60-s segments (Bertucci et al. 2020). Reprinted by permission from Springer Nature. Bertucci F, Guerra AS, Sturny V, et al., A preliminary acoustic evaluation of three sites in the lagoon of Bora Bora, French Polynesia. Environ Biol Fishes 103:891–902; https://doi.org/10.1007/s10641-020-01000-8. © Springer Nature, 2020. All rights reserved

soundscape of a shallow-water Marine Protected Area in the Mediterranean Sea. Seasonal patterns were seen within the lower (63 Hz–1 kHz) and higher (4–64 kHz) OBLs due to increases in wind in winter and snapping shrimp activity in summer, respectively. In contrast, sound levels within the 2-kHz octave band remained stable as sound from both wind and snapping shrimp entered this frequency band, thus attenuating seasonal fluctuations. Sound levels in the 1/3 octave bands centered at 63 and 125 Hz were set as indicators of ship noise by the European Commission Joint Research Centre (Tasker et al. 2010). Ship noise studies in shallow water, however, highlight that natural sound sources (i.e., wind) and propagation characteristics may render these indicators less useful in coastal areas and that bandlevels at 200 and 315 Hz should be included, particularly in areas frequented by smaller recreational vessels (Garrett et al. 2016; Picciulin et al. 2016).

### 7.5.2 Identification of Sound Sources

Soundscape ecology involves the identification of sound sources and whether they are part of the biophony, geophony, or anthropophony. Most sources have a unique sound signature (see examples earlier in this chapter), which can be identified from power spectra. Knowing to which soundscape component a sound belongs helps to evaluate how pristine an environment is and pinpoint possible impacts from human activities. Choruses by insects (Brown et al. 2019), anurans (Nityananda and Bee 2011), birds (Baker 2009), marine invertebrates (Radford et al. 2008), and fish (Parsons et al. 2016) are so distinct that they are easily identified as biophony. Knowledge on species-specific vocalizations helps to monitor species behavior and species-specific responses to environmental stressors (such as noise) as demonstrated with insects (e.g., Walker and Cade 2003), amphibians (e.g., Gibbs and Breisch

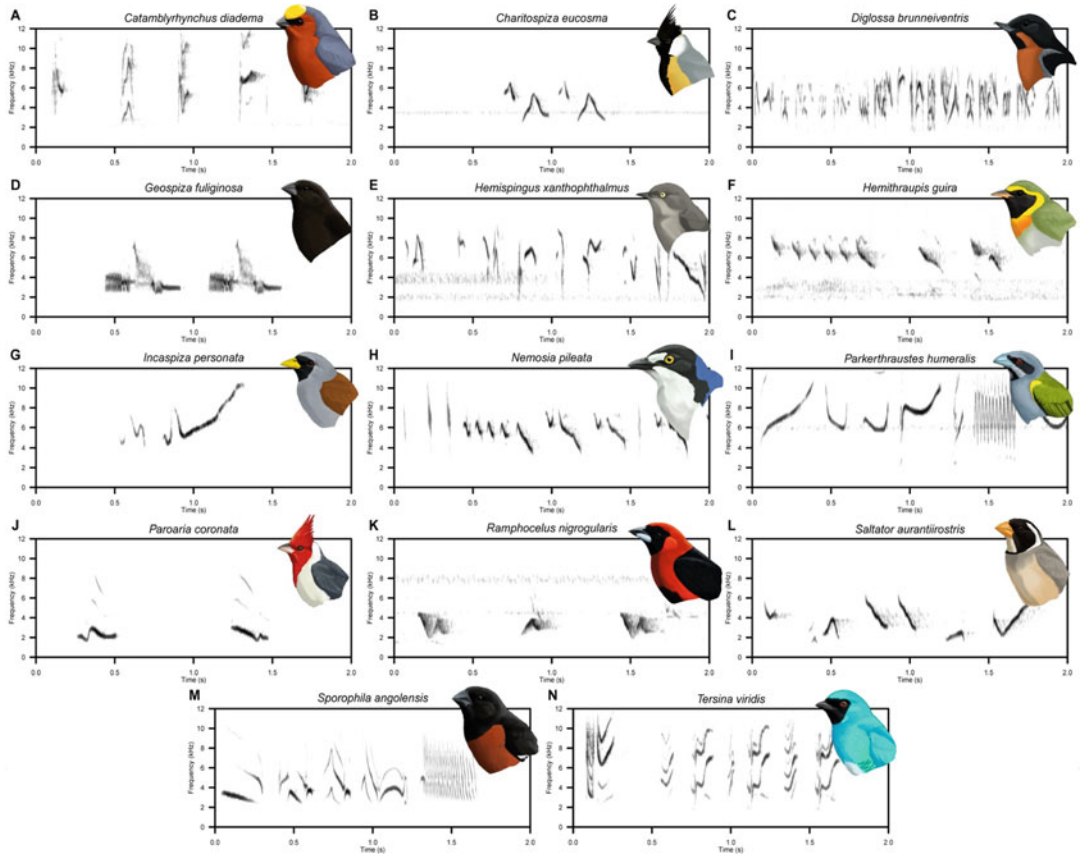



**Fig. 7.22** Spectrograms highlighting the difference in vocalizations between 14 different tanager species, which can be used to monitor behavior and response to environmental change (Mason and Burns 2015). Reprinted by permission from Oxford University Press. Mason NA, Burns KJ, The effect of habitat and body size on the evolution of vocal displays in Traupidae (tanagers), the largest family of songbirds. Biol J Linn Soc 114:538–551; https://doi.org/10.1111/bij.12455. © The Linnean Society of London, 2015; https://global.oup.com/academic/rights/permissions/. All rights reserved. Reuse requires permission from OUP

2001), birds (Fig. 7.22; e.g., Jahn et al. 2017), and mammals (e.g., Nijman 2001; Parks et al. 2007). Similarly, the sounds of the geophony and anthropophony have characteristic spectral features by which they can be identified.

Studies differ, however, in their methodology to identify sound sources. By listening to sounds while observing their spectrograms in real-time (see Sect. 7.5.3.1), experts can employ their personal experience to separate biotic and abiotic sounds and to identify species. Alternatively, sounds can be compared to labeled recordings in sound libraries (see URLs at the end of this chapter) and spectrograms can be compared to those found in the literature. However, manual inspection of sound files is labor intensive; and so, some studies make use of automatic detection and classification software (see Chap. 8).

## 7.5.3 Visual Displays of Soundscapes

### 7.5.3.1 Spectrograms

A spectrogram displays acoustic power density as a function of time and frequency. Each column in the spectrogram is a result of Fourier-transforming a section of the recorded time series of sound pressure. The frequency and time resolutions of the spectrogram are affected by

the window length and type of window function used (see Chap. 4). Techniques such as zero-padding (i.e., expanding a time window with zeros) and overlapping time windows may enhance the apparent resolution in frequency and time. Each pixel (or cell) of the spectrogram eventually represents an average sound power, averaged into time and frequency bins. Spectrograms are a useful tool to examine the time, frequency, and amplitude details of a sound at different time scales, potentially identifying the sound source. Spectrograms that contain the vocalizations of multiple sound sources can provide information on species vocal dynamics, acoustic niches, and how animals may be affected by acoustic changes in their surroundings. For example, mixed anuran species' breeding choruses in Minnesota, USA, revealed acoustic niche partitioning within the frequency domain (Fig. 7.23), while fin whale vocalizations were masked by ship noise in Italy (Fig. 7.24).

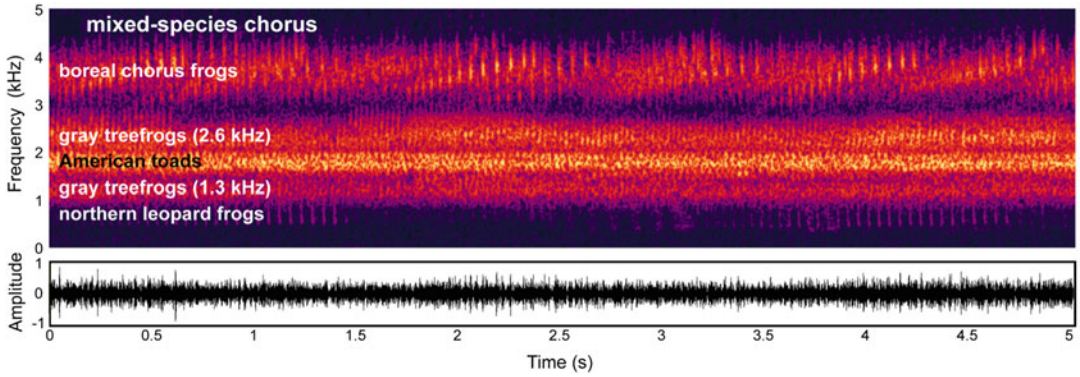



**Fig. 7.23** Anuran choruses recorded in Minnesota comprising calls of four species. Note the occupation of different frequency bands by these species, suggesting acoustic niche partitioning within the frequency domain. Modified image; © Nityananda and Bee (2011); https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0021191. Published under CC BY 4.0; https://creativecommons.org/licenses/by/4.0/

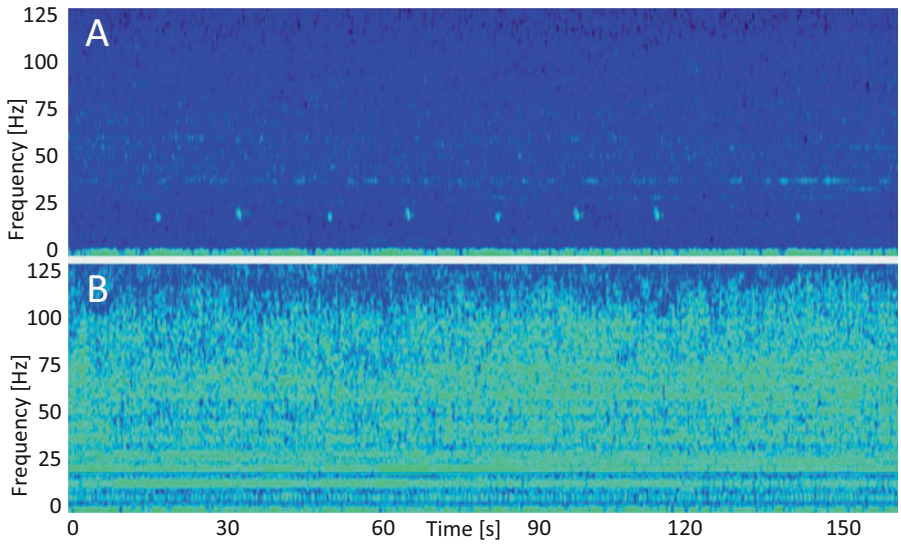



**Fig. 7.24** Spectrograms of (**a**) 20-Hz fin whale vocalizations off Sicily, Italy, and (**b**) a passing ship, which masked the fin whale sounds

Long-term monitoring programs typically make use of long-term spectral averages (LTSAs), which are spectrograms that were averaged into observation windows much longer than the underlying FFT windows. Observation windows may range from tens of seconds, to one minute, to several hours, to the length of one recording within a duty cycle (e.g., Gavrilov and Parsons 2014). LTSAs highlight persistent soundscape contributors (e.g., shipping or storms), repetitive soundscape contributors (e.g., night-time choruses), and dominant events (e.g., an earthquake). They can be used to identify specific days or hours rich in sounds, quiet versus noisy periods, or correlations between acoustic patterns and environmental factors. Fig. 7.25 shows a 3-week LTSA, in which dominant events were marked (e.g., nightly fish chorus, whale choruses, stormy days, and passing ships). Break-out spectrograms show specific signals on a finer temporal scale (Erbe et al. 2016b). Alternatively, long-term spectrograms may display minimum (LTSmin), maximum (LTSmax), median (LTSmed), or other percentile levels (e.g., LTS75), computed within each frequency bin over some time window (Righini and Pavan 2020). The minima will track the quietest baseline and the maxima can highlight strong but brief

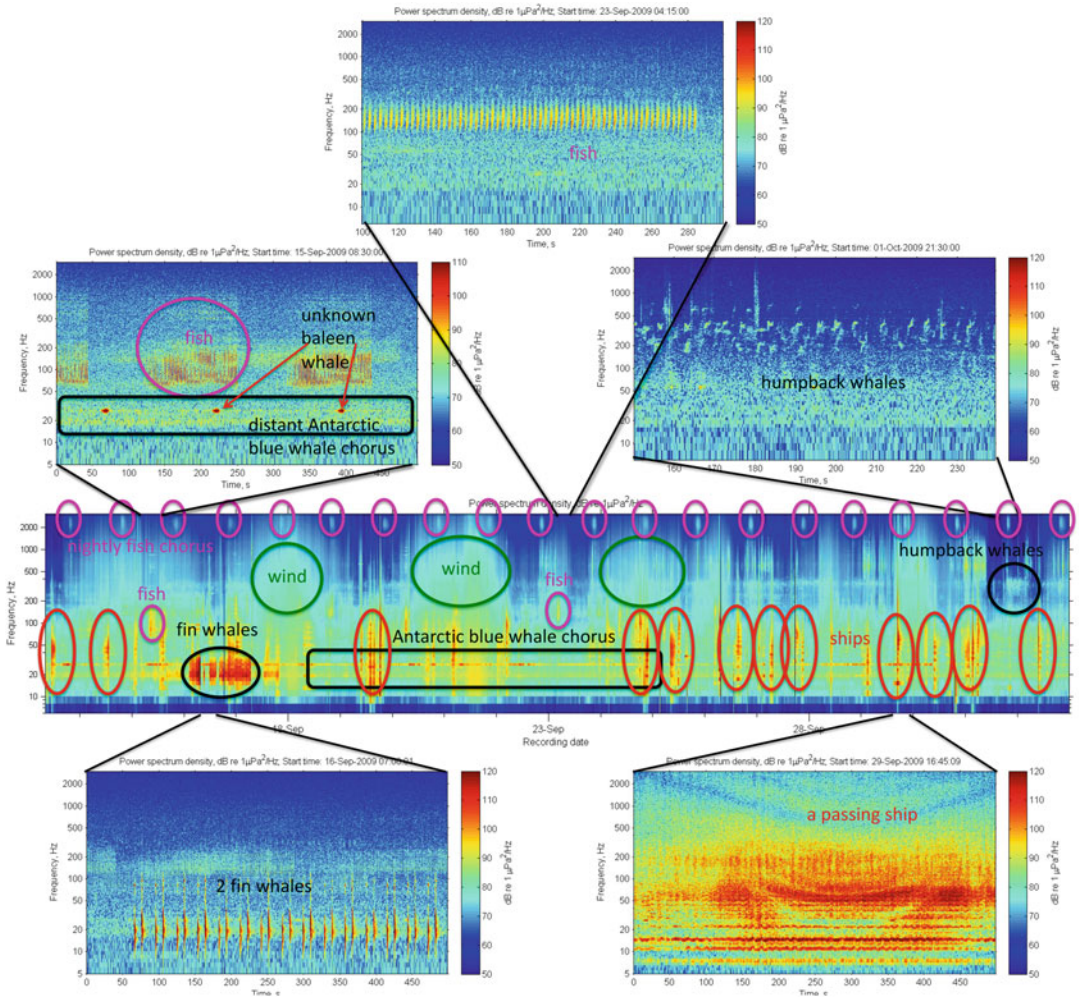



**Fig. 7.25** Spectrograms of the marine soundscape in the Perth Canyon, Australia. Middle panel shows a 3-week LTSA, computed with a 10-min observation window. The surrounding panels display short-term spectrograms of example sounds (Erbe et al. 2016b)

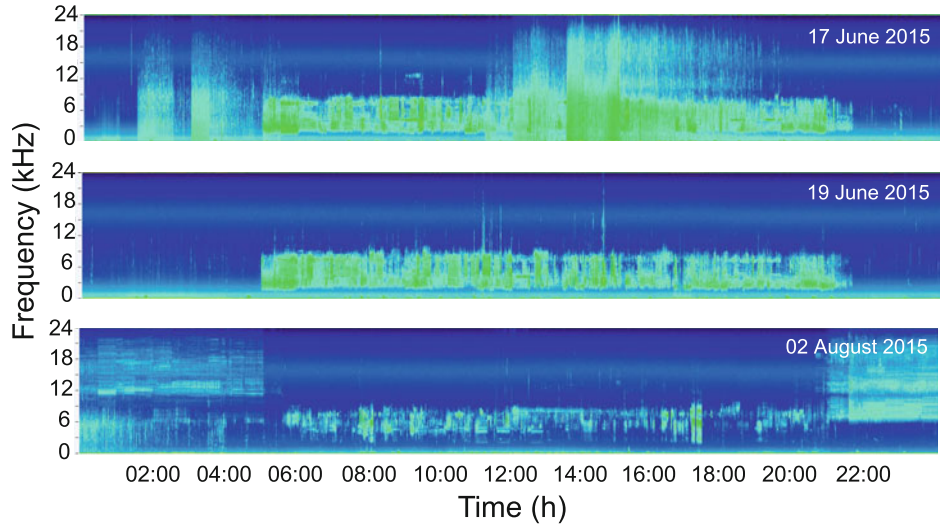



**Fig. 7.26** LTSmax spectrograms from the same location (Sasso Fratino Integral Nature Reserve, Italy) on three different dates and under different weather conditions. Biophony is concentrated between 1.5 and 9 kHz and decreased in August. LTSmax produced with *SeaPro* software by combining 48 frames of 10 min each, recorded every 30 min (Righini and Pavan 2020)

events, which would otherwise be averaged and potentially missed in LTSAs. Fig. 7.26 shows three 24-h LTSmax of an Italian soundscape on different dates and under different weather conditions (Righini and Pavan 2020). The images show sound sources present from midnight to midnight: (top) one day in June 2015 with some bursts of rain, (middle) one day with good weather and a clear image of the biophony concentrated between dawn and dusk in the frequency range 1.5–9 kHz, and (bottom) one day recorded in August, with a less dense biophony during daylight hours but Orthopteran choruses in the night. In August, a short period of light rain is also shown on the left side. In addition, the stream noise below 1 kHz in August was lower than in June. The faint band between 12 and 18 kHz present in all 3 panels was due to the intrinsic noise of the recorder.

#### 7.5.3.2 Power Spectral Density Percentile Plots

While spectrograms (including LTSAs) show how the sound spectrum changes over time (from one FFT window to the next or from one LTSA observation window to the next), there might be a need to quantify this variability. Power spectral density (PSD) percentile plots quantify the spectrum variability over the duration of a temporal analysis window. PSD is plotted against frequency. At each frequency, several percentile levels are shown, commonly the median (50th percentile) and the quartiles (25th and 75th percentiles), but perhaps also additional percentiles (e.g., 1st, 5th, 95th, and 99th). The nth percentile gives the levels that were exceeded n% of the time. There is no standard for the length of the temporal analysis window, and selection depends on the specific study questions. Temporal analysis windows of 24 h, one season, or one full year are common. Dominant contributors to the soundscape can then be identified by the shape and levels of the curves. Additional information is provided by plotting the Spectral Probability Density (SPD) as background colors that represent the probability of levels being reached based on normalized histograms of sound levels within each frequency bin (Fig. 7.27; Merchant et al. 2013). Merchant et al. (2015) gave detailed information on how to compute PSDs and SPDs with their publicly available software *PAMGuide*. Also see Chap. 4.

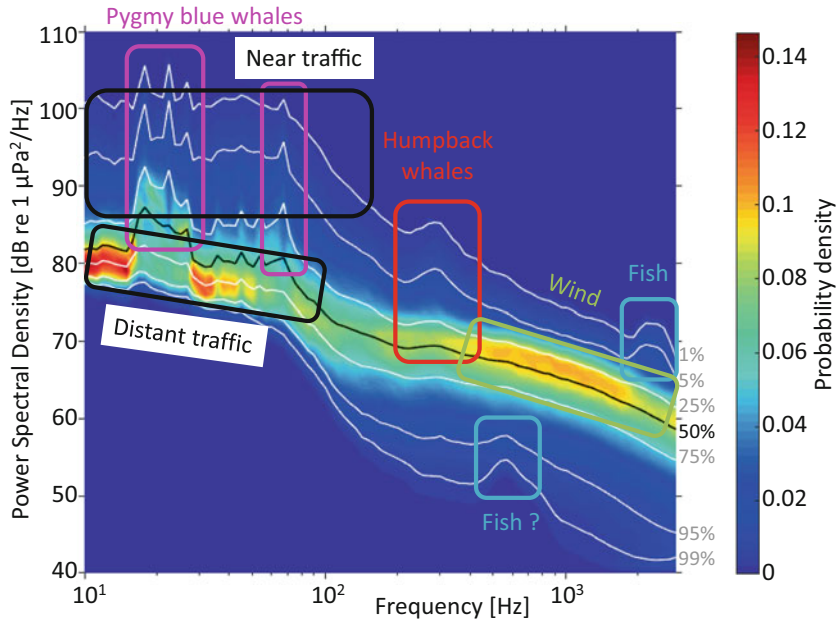



**Fig. 7.27** Plot of power spectral density percentiles and probability density for the annual soundscape of the Perth Canyon, Australia. The strongest sound sources were pygmy blue whales and nearby ships at 10–200 Hz, humpback whales at 300 Hz, and fishes at 2 kHz, whereas the most common sound sources were distant shipping at 10–100 Hz and wind at 300 Hz–3 kHz (Erbe et al. 2016b)

#### 7.5.3.3 Soundscape Maps

Soundscape maps literally show sound levels on a map. Such maps are mostly produced by modeling sound propagation from multiple sources, distributed over the area. Model results may be validated by point measurements (i.e., recordings at selected places; Erbe et al. 2014, 2021; Schoeman et al. 2022). Sound maps may be produced for specific frequencies of interest (e.g., relevant to human audiology; Bozkurt and Demirkale 2017) or for a specified receiver height or depth (e.g., migrating whales below the sea surface; Tennessen and Parks 2016; Bagočius and Narščius 2018). Sound propagation maps typically focus on specific sound sources (e.g., highways or railways; Fig. 7.28; Aletta and Kang 2015; Drozdova et al. 2019).

Maglio et al. (2015) developed a near real-time model that shows the propagation of sound from individual ships in the Ligurian Sea. However, focus can also be placed on cumulative or average sound levels over a specified time frame to identify areas of long-term risk to humans or animals from noise exposure. Erbe et al. (2012) computed a map of average sound levels from annual ship tracks to highlight areas along the Canadian coast where ship noise exceeded the European criterion of 100 dB re 1 µPa rms (Fig. 7.29). The same concept was later used to identify areas where (a) strong sound levels overlapped with high animal density (identifying areas of risk; Fig. 7.30; Erbe et al. 2014), and (b) low sound levels overlapped with high animal density (identifying areas of opportunity for conservation management; Fig. 7.30; Williams et al. 2015).

### 7.5.4 Acoustic Indices

Apart from sound level statistics (such as SPL measures, PSD percentiles, and SPD), additional metrics, such as acoustic indices, exist, which may quantify soundscapes as a whole or quantify

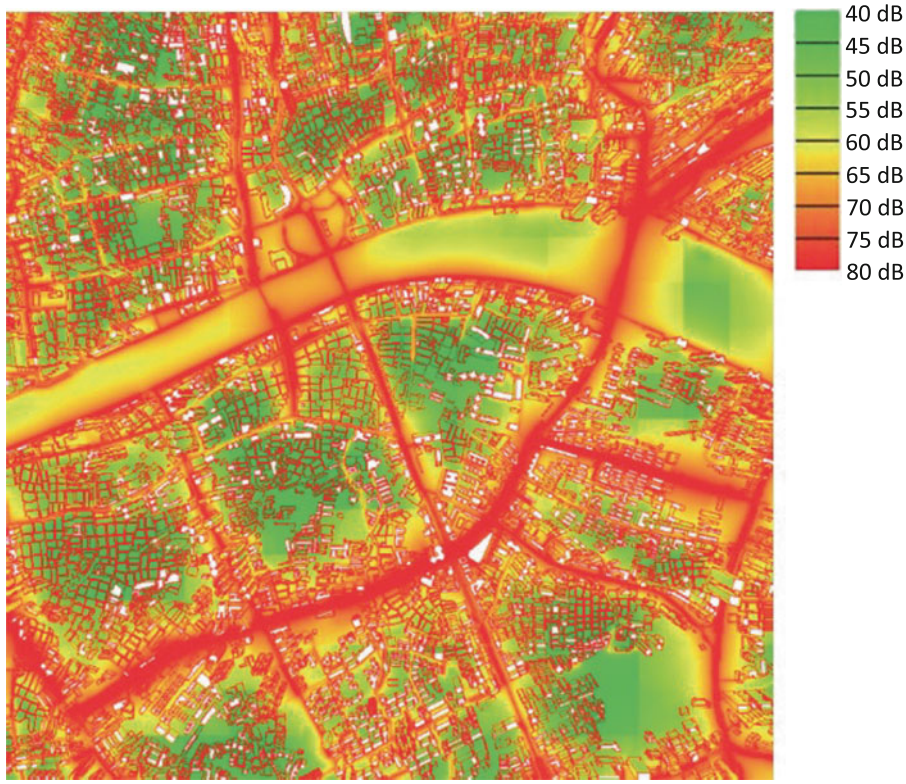



**Fig. 7.28** Noise-map of a roadway in an urban area. Red indicates highest noise levels and green represents the quietest areas. © Cai et al. 2018; https://www.hindawi.com/journals/jat/2018/7031418/fig4/. Published under CC BY 4.0; https://creativecommons.org/licenses/by/4.0/

the biophony, geophony, and anthropophony separately or in comparison. Acoustic indices can be used as a tool to assess the quality of soundscapes and the underlying ecosystem. Historically, researchers assessed the number of species (i.e., species richness) and number of individuals belonging to each species (i.e., species evenness) by counting the number of acoustic identifications while walking along survey transects or listening to recordings (Obrist et al. 2010). However, this approach is inefficient, subjective, and limited to brief observation times. In contrast, a transect or grid of automated recording systems allows acoustic surveys in remote areas, over extended periods, and in most field conditions (Acevedo and Villanueva-Rivera 2006).

To support the analyses and interpretation of consequent large datasets, researchers have been developing acoustic indices that summarize and score the structure and distribution of acoustic power over frequency and/or time, reflecting a correlation with species presence and distribution (e.g., Towsey et al. 2014). While traditionally developed for terrestrial communities, acoustic indices are now also increasingly applied to the aquatic environment (e.g., Parks et al. 2014; Harris et al. 2016; Bolgan et al. 2018a). In particular when the same instruments and protocols are used, acoustic indices allow for comparisons of soundscapes between multiple sites recorded over the same period or an evaluation of the changes of a soundscape over time (Righini and Pavan 2020; Farina et al. 2021).

Examples of acoustic indices include:

1. Bioacoustic Index (BI): Aims to quantify biophonic activity by thresholding spectral power in biophony-specific frequency bands (Fig. 7.31; Boelman et al. 2007),

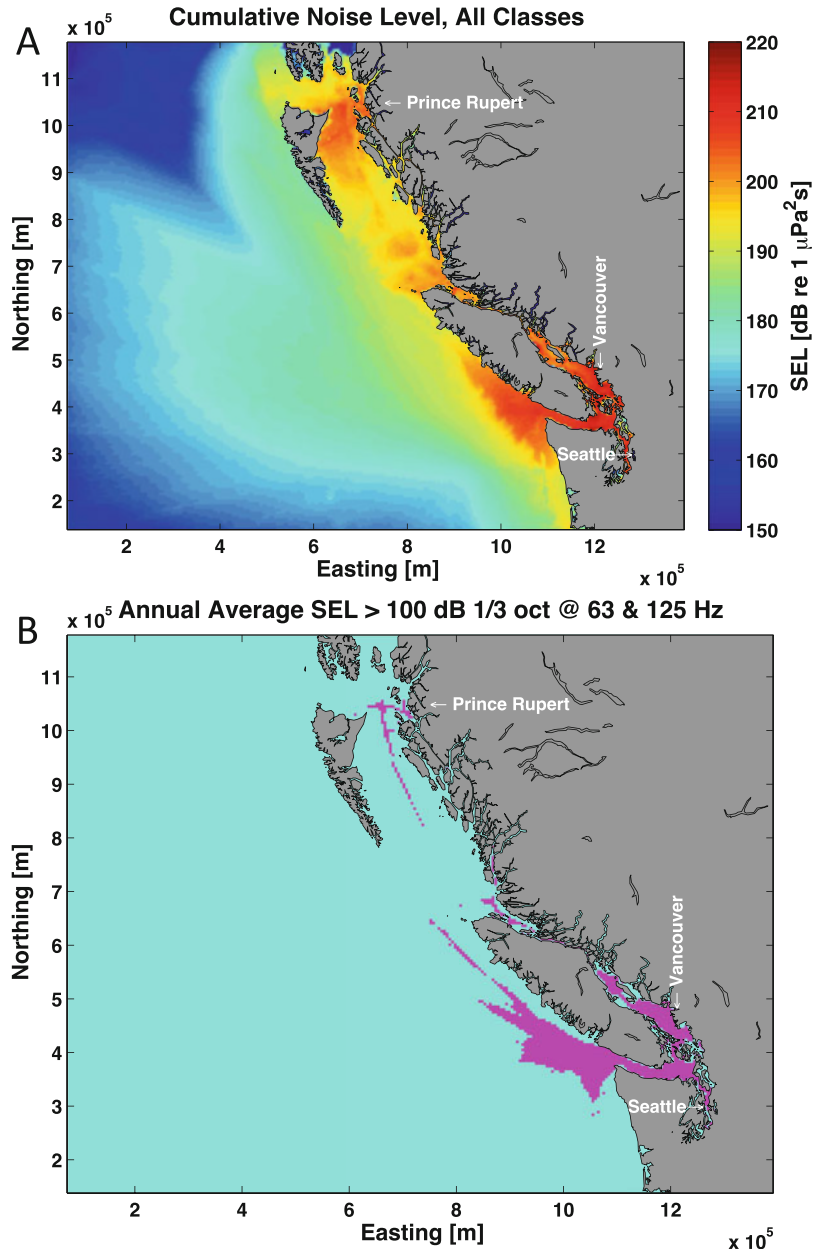



**Fig. 7.29** Illustration of the conversion of cumulative hours of ship traffic along the Canadian coast to cumulative noise levels (**a**) to identify areas where annual average received levels exceeded the European criterion for low-frequency ambient noise of 100 dB re 1 μPa rms (**b**; Erbe et al. 2012). © Acoustical Society of America 2012. All rights reserved

2. Entropy Index (H): Equals the product of two sub-indices, spectral (Hf) and temporal entropy (Ht), computed on the average frequency spectrum and on the Hilbert amplitude envelope of the raw bioacoustic signal, respectively (Sueur et al. 2008b),
3. Acoustic Diversity Index (ADI): Divides the spectrum into specific frequency bins, selects the bins surpassing a preset power threshold, and applies the Shannon entropy to these bins (Villanueva-Rivera et al. 2011),
4. Acoustic Evenness Index (AEI): Divides the spectrum into specific frequency bins, selects the bins surpassing a preset power threshold, and considers the distribution of strong frequency bins by computing the Gini coefficient (Villanueva-Rivera et al. 2011),

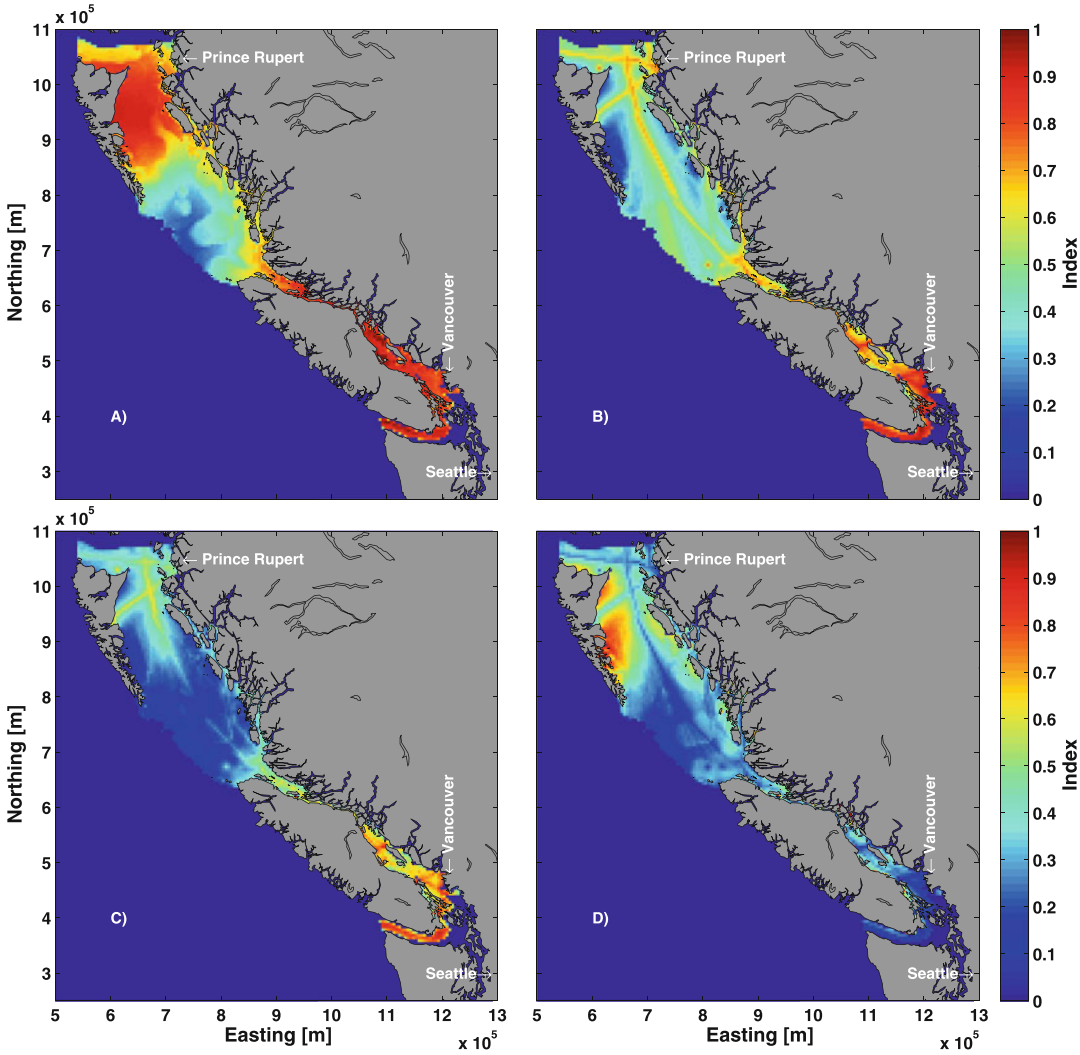



**Fig. 7.30** Maps of (**a**) harbor porpoise (*Phocoena phocoena*) density, (**b**) audiogram-weighted ship noise, (**c**) areas of risk (i.e., high animal density and high noise), and (**d**) areas of opportunity (i.e., high animal density and low noise) in British Columbia, Canada. © Williams et al. 2015; https://doi.org/10.1016/j.marpolbul.2015.09.012. Licensed under CC BY-NC-ND 4.0; https://creativecommons.org/licenses/by-nc-nd/4.0/

5. Acoustic Complexity Index (ACI): Measures the temporal variation in acoustic power by calculating sequential power differences (from one FFT window to the next), in all frequency bands separately, then sums over frequency (Fig. 7.31; Pieretti et al. 2011), and
6. Normalized Difference Soundscape Index (NDSI): Equals the ratio of low-frequency (indicative of anthropophony) to high-frequency power (indicative of biophony) to capture the level of anthropogenic disturbance (Kasten et al. 2012).

These and other indices are coded in shareware *R* packages, such as *seewave* (Sueur et al. 2008a; Sueur 2018), *soundecology* (Villanueva-Rivera and Pijanowski 2018), and *bioacoustics* (Marchal et al. 2020). However, the analysis of long-term recordings can also aim at recognizing individual

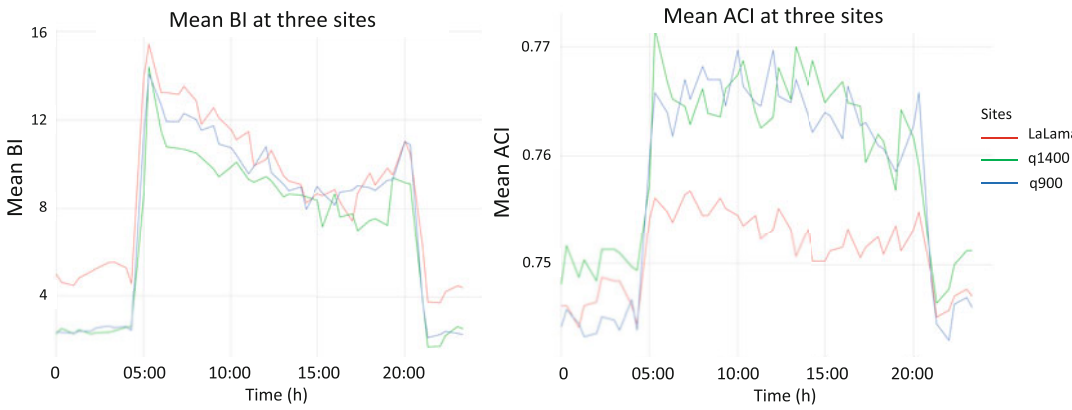



**Fig. 7.31** Bioacoustic Index (BI) and Acoustic Complexity Index (ACI) for three Italian locations in the Integral Nature Reserve of Sasso Fratino, Italy, showing a strong peak at sunrise, followed by a gradual decline with a second peak at sunset

species' signatures by listening, by observing spectrograms, and by using sound recognition tools to identify the presence and recurrence of defined sound models. The *R* package *monitoR* (Katz et al. 2016) can be used to identify user-defined sound models.

It should be noted that acoustic indices applied in two different environments can produce confounding results and so the robustness of these indices to environmental change and to different soundscape compositions has been questioned (Harris et al. 2016; Bolgan et al. 2018a).

Parks et al. (2014) found that seismic airgun pulses interfered with the Entropy Index and therefore did not accurately reflect species richness within the Atlantic Ocean where seismic surveys were commonly detected. Bolgan et al. (2018a) assessed the robustness of the Acoustic Complexity Index to fine variations in fish sound abundance (i.e., number of sounds) and diversity (i.e., number of different calls); both changed index values. Hence, it would be difficult to infer whether a change in this index resulted from a change in fish abundance or fish species diversity. Biophony and anthropophony can overlap in frequency and time as well as vary with frequency and time. Acoustic index performance depends greatly on the frequency and time resolutions used in the computation of the various quantities and is affected by temporal (and spatial) patterns as well as local (and temporally variable) sound propagation conditions (Mooney et al. 2020). As a result, acoustic indices are sometimes tuned for specific environments, limiting comparability across environments and time.

## 7.6 Applications of Soundscape Studies

Soundscape studies can reveal information on animal distribution, abundance, and behavior; species diversity; and changes of all of these over time under environmental and human influences. Hence, soundscape analyses can be used as ecological tools to understand, conserve, and restore soundscapes as part of conservation management plans (Pavan 2017).

### 7.6.1 Conservation of Natural Soundscapes

#### 7.6.1.1 Management

Documenting, analyzing, and understanding a soundscape can provide important information for wildlife and habitat managers on species

richness, animal behavior patterns, effects of anthropogenic sounds, land-use, and climate change. Documenting relatively pristine soundscapes before they disappear (Righini and Pavan 2020; Farina et al. 2021) can aid re-establishment of degraded acoustic habitats through habitat restoration, animal relocation, elimination of invasive species, or restrictions of activities that generate anthropogenic sound and affect animal behavior. The success of soundscape restoration can then be demonstrated through acoustic monitoring and analysis (Pavan 2017).

Development and implementation of a comprehensive acoustic monitoring program can aid management of a protected area in several ways. Firstly, storage of quantitative data about the acoustic environment can be used to create pivotal repositories for immediate or future analyses of spatial and temporal patterns and differences at large scales. LTSA spectrograms, for example, provide a summary of day-by-day acoustic settings and the possibility to display information, not only on the diversity of acoustic species (as in a census) but also on the density and richness of the biophonic components. The study of an Integral Nature Reserve (Sasso Fratino, Casentinesi Forests National Park, Italy) demonstrated that the biophony dominated both geophony and anthropophony, with undisturbed daily cycles (Righini and Pavan 2020; Farina et al. 2021). Secondly, monitoring soundscapes can help managers detect unwanted and unlawful activities in protected areas. Human voices can be used to identify trespassers, gunshots to locate hunters and poachers, humming chainsaws to find illegal logging, vehicle sounds to document unauthorized vehicle use, and sounds from livestock to pinpoint unlawful grazing. Wrege et al. (2017) found that gunshot sounds within a closed-canopy forest of the Congo could be detected over a 7–10 $km^2$ area, depending on the gun used and orientation to the acoustic receiver. Eight years of acoustic monitoring did not reveal a correlation between illegal hunting of forest elephants (*Loxodonta cyclotis*) and time of day or season. However, hunting intensity seemingly decreased after initiating patrols in 2009, highlighting the potential use of soundscape studies to monitor for illegal human activities and to assess the effectiveness of conservation efforts.

Investigation of underwater soundscapes can also aid in the detection of foreign vessels by the military, unauthorized commercial fishing vessels, unlawful vessels in restricted areas (i.e., no-go zones or marine protected areas; Kline et al. 2020), and illegal fishing activities with explosives (Xu et al. 2020).

#### 7.6.1.2 Education

The rates of biodiversity loss, habitat loss, invasion of alien species, and species extinctions are high (Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services [IPBES] 2019). Helping citizens and stakeholders appreciate biodiversity is a necessity to establish a general willingness to address anthropogenic causes of ecosystem demise. In this context, animal sound and soundscape recordings not only serve science but have the potential to trigger people's curiosity to learn more about the importance of ecosystems and their preservation, which will lead to conservation efforts. Such transfer of science, via education, to conservation has been demonstrated in several case studies (e.g., Padua 1994; Macharia et al. 2010; Pavan 2017; Barthel et al. 2018). Exhibits and educational programs on the sounds from nature in museums, zoos, park visitor centers, and websites can stimulate interest in and care about the acoustic environment. An example is Bernie Krause's *Great Animal Orchestra* exhibition[1]. Alternatively, listening to animal sounds during a guided nature walk can generate an appreciation for soniferous animals, which can result in long-term public engagement and commitment to conservation by citizen scientists. Soundscape studies can help to create publicly available sound libraries and help to identify areas within a park for visitors to experience songbirds, calling frogs, chorusing insects, waterfalls, rushing streams, etc. One example of integrating soundscape monitoring and education is the Natural Sound Program, established in

[1] https://thevinylfactory.com/features/bernie-krause-great-animal-orchestra/; accessed 27 September 2020

2000 by the U.S. National Park Service (National Park Service [NPS] 2000). This program aims to manage the acoustic environment while providing for educational and inspirational visitor experiences.

### 7.6.2 Monitoring the Health of Agroecosystems

High productivity from agricultural fields can be maintained through insecticides, pesticides, and fertilizers, but the use of these products may result in chemical pollution with consequent loss of plant and animal biodiversity (e.g., Carson 1962; Boatman et al. 2004; Kerr and Cihlar 2004; Kleijn et al. 2009). Hence, habitats connected to agricultural lands might exhibit poorer soundscapes. In contrast, organic farmers strive to maintain productivity through natural agroecosystems, ensuring environment quality and ecological balances. Bird, insect, amphibian, and bat communities serve as indicators of ecosystem health, and an agroecosystem should have a balance of mixed species that provide natural pest control. The ecological quality of an agroecosystem can therefore be evaluated by the species-richness of its soundscape (e.g., Hole et al. 2005; Kleijn et al. 2011; Pavan 2017). Doohan et al. (2019) identified bird and bat species-specific or guild-specific bioindicators as successful biomonitoring tools for agricultural industries. Systematic monitoring of biological sounds can provide an accurate and practical assessment tool for farmers, policymakers, researchers, and others interested in maintaining or restoring farmland ecosystems, and ultimately encourage the adoption of beneficial and sustainable farming practices.

### 7.6.3 Improving Captive Animal Welfare

Noise may be omnipresent for captive animals in livestock-operations, zoos, aquaculture, and aquaria. While wind and rain contribute naturally to ambient sound in outdoor animal enclosures (Wiseman et al. 2014), anthropogenic sound from mechanical devices (e.g., Wysocki et al. 2007; Scheifele et al. 2012b), background music (Scheifele et al. 2012a), and visitors (e.g., Quadros et al. 2014; Sherwen and Hemsworth 2019) is characteristic of many indoor, outdoor, and underwater animal holding facilities. O'Neal (1998), for example, found that underwater sound pressure levels were 25 dB (20–6400 Hz) louder in exhibits inside the Monterey Bay Aquarium than in a nearby natural offshore environment, predominantly due to sound from machinery. Similarly, Scheifele et al. (2012b) detected an increase in sound pressure levels by 10–20 dB (20 Hz–1 kHz) when air pumps were switched on within the Georgia Aquarium. These increases in sound levels can have adverse effects on animal welfare because of physiological and behavioral changes (e.g., Owen et al. 2004).

Sound sources that may impact animals might not be audible to humans, and so animal keepers might not be aware of acoustic disturbance to kept animals. For example, laboratory mice are sensitive to ultrasound, above the human hearing range. Laboratory equipment (e.g., air conditioners and lighting) may emit ultrasound and, unknown to humans, stress animals within these facilities (Sales et al. 1988). Identifying such sources is necessary for the improvement of acoustic conditions to increase captive animal welfare (De Queiroz 2018). Sound can further be exacerbated by hard reflective surfaces and the geometry of an exhibit; hence, some noise problems can be solved by improving exhibit design (Wark 2015; De Queiroz 2018). Restricting visitor group sizes, reducing operation hours, limiting the number of shows, and reducing the level of background music can also mitigate negative impacts of noise on captive animals.

## 7.7 Conclusion

Soundscapes are composed of a myriad of sounds that can be grouped into biophony, geophony, and anthropophony based on their origin. Natural soundscapes have ecological value and modifying these natural assets could

lead to changes in ecosystem functioning and biodiversity. At present, natural soundscapes are disappearing at an unprecedented rate because of human interference. Human activities create sound, change land-use patterns, directly remove animals from their habitat through overharvesting and illegal hunting, and lead to climate change, thereby directly and indirectly affecting both geophony and biophony. Soundscape studies can be used as an ecological tool to study animal distribution, behavior, biodiversity, and the effects of environmental stressors (such as anthropogenic noise or climate change). Soundscape studies can subsequently inform conservation management and assess the effectiveness of management and conservation efforts.

## 7.8 Additional Resources

Below is a selection of free, online resources; last accessed 20 June 2022.

### 7.8.1 Sound Libraries

Sound libraries can serve as reference during the identification of sound sources. They are also an educational tool to create awareness of the myriad of sounds that may contribute to a soundscape.

- The Macauley library from the Cornell Lab of Ornithology contains a large collection of biophony: https://search.macaulaylibrary.org/catalog?view=List&searchField=animals
- The Discovery Of Sound In The Sea (DOSITS) website, developed by the University of Rhode Island Graduate School of Oceanography in partnership with Marine Acoustics Inc., contains an underwater sound library as well as a collection of easy-to-read scientific information on sound in the ocean: https://dosits.org
- The sounds of Australian and Antarctic marine mammals, Curtin University: https://cmst.curtin.edu.au/research/marine-mammal-bioacoustics/
- A collection of biophony, geophony, and various soundscape recordings from all over the world, the British Library: https://sounds.bl.uk/Environment
- Sounds recorded by National Park Service researchers in U.S. National Parks, such as Yellowstone National Park and Rocky Mountain National Park: https://www.nps.gov/subjects/sound/gallery.htm
- A collection of biophony (i.e., invertebrates, amphibians, fishes, reptiles, birds, and mammals), Museum für Naturkunde. Note that some sound descriptions are in German: https://www.museumfuernaturkunde.berlin/en/science/animal-sound-archive
- A collection of biophony, SeaWorld Parks and Entertainment: https://seaworld.org/animals/sounds/
- A collection of marine biophony, geophony, and anthropophony, Ocean Conservation Research: https://ocr.org/sound-library/
- The Xeno-Canto collection of animal recordings provided by scientists and amateur recordists: https://www.xeno-canto.org/
- Web pages of the University of Pavia about bioacoustics and ecoacoustics, including samples of sounds: http://www.unipv.it/cibra

### 7.8.2 Ocean Acoustic Observatories

Ocean acoustic observatories provide a continuous stream of acoustic data either in real-time or archived:

- Australia's Integrated Marine Observing System (IMOS): https://imos.org.au/facilities/nationalmooringnetwork/acousticobservatories
- Indian Ocean Acoustic Observatory OHASISBIO: https://www-iuem.univ-brest.fr/lgo/les-chantiers/ohasisbio/?lang=en
- Listening to the Deep Ocean (LIDO): http://www.listentothedeep.net/
- Monterey Bay Aquarium Research Institute (MBARI): https://www.mbari.org/soundscape-listening-room/

### 7.8.3 Software for Soundscape Analysis

- *Characterization Of Recorded Underwater Sound (CHORUS)*, a *MATLAB* (The MathWorks Inc., Natick, MA, USA) graphic user interface developed by Curtin University: https://cmst.curtin.edu.au/products/chorus-software/ (Gavrilov and Parsons 2014).
- *PAMGuard* for passive acoustic monitoring: http://www.pamguard.org/download.php?id=108
- *Triton Software Package*, a *MATLAB* graphic user interface developed at Scripps Institution of Oceanography: http://www.cetus.ucsd.edu/technologies_triton.html
- *OSPREY*, a *MATLAB* graphic user interface developed by Oregon State University: https://www.mobysound.org/software.html
- *R* package *seewave* available for download from within *RStudio*: https://cran.r-project.org/web/packages/seewave/index.html
- *R* package *soundecology* available for download from within *RStudio*: https://cran.r-project.org/web/packages/soundecology/index.html
- *R* package *bioacoustics* available for download from within *RStudio*: https://cran.r-project.org/web/packages/bioacoustics/index.html
- *SoundRuler* for measuring acoustic signals: http://soundruler.sourceforge.net/main/
- *Sound Analysis Pro* for analysis of biophony: http://soundanalysispro.com
- *SeaPro* and *SeaWave* for recording, analysis, and real-time display of bioacoustic signals and biophony: http://www.unipv.it/cibra/seapro.html
- *SOX* a command line tool for sound file manipulation and analysis: https://sourceforge.net/projects/sox/
- *Raven Lite* to record, save, and visualize sounds as spectrograms and waveforms: https://ravensoundsoftware.com/software/raven-lite/

### 7.8.4 Software for Sound Propagation Modeling

- The *Acoustic Toolbox User interface and Post processor* (*AcTUP*) written in *MATLAB* for modeling range-independent and range-dependent environments: http://cmst.curtin.edu.au/products/underwater/ (Duncan and Maggi 2006).
- Graphical user interface *i-Simpa* suitable for 3D indoor sound propagation modeling as well as for modeling of environmental noise: https://i-simpa.ifsttar.fr/download/download0/
- Software tool created by the *openPSTD* project to aid sound propagation modeling in urban environments: http://www.openpstd.org/Download%20openPSTD.html
- The *NoiseModelling* tool designed to create environmental noise maps of large urban areas: https://noise-planet.org/noisemodelling.html
- The ArcGIS toolbox *SPreAD-GIS* for modeling engine noise propagation in natural areas incorporating atmospheric, wind, vegetation, and terrain effects (Reed et al. 2010).

### 7.8.5 Software for Automatic Signal Detection

Some of the software packages for soundscape analysis include signal detectors:

- *CHORUS* includes detectors for pygmy blue whale song, fin whale 20-Hz downsweeps, and an unidentified spot-call.
- *PAMGuard* includes detectors for odontocete and mysticete vocalizations.

Other automatic signal detection resources:

- *R* package *monitoR* available for download from: https://cran.r-project.org/web/packages/monitoR/index.html
- *Ishmael*: http://bioacoustics.us/ishmael.html

## References


Acevedo MA, Villanueva-Rivera LJ (2006) Using digital recording systems as effective tools for the monitoring of birds and amphibians. Wildl Soc Bull 34:211–214. https://doi.org/10.2193/0091-7648(2006)34[211:UADRSA]2.0.CO;2

Ainslie M (2010) Principles of sonar performance modelling. Springer, Heidelberg

Aletta F, Kang J (2015) Soundscape approach integrating noise mapping techniques: a case study in Brighton, UK. Noise Mapp 1:1–12. https://doi.org/10.1515/noise-2015-0001

Amorim MCP, Vasconcelos RO, Fonseca PJ (2015) Fish sounds and mate choice. In: Ladich F (ed) Sound communication in fishes. Springer, Vienna, pp 1–33

Anderson RC, Herrera M, Ilangakoon AD et al (2020) Cetacean bycatch in Indian Ocean tuna gillnet fisheries. Endanger Species Res 41:39–53. https://doi.org/10.3354/esr01008

Andrew RK, Howe BM, Mercer JA (2011) Long-term trends in ship traffic noise for four sites off the north American west coast. J Acoust Soc Am 129:642–651. https://doi.org/10.1121/1.3518770

Arch VS, Grafe TU, Narins PM (2008) Ultrasonic signalling by a Bornean frog. Biol Lett 4:19–22. https://doi.org/10.1098/rsbl.2007.0494

Attenborough K, Bashir I, Hill TJ, Taherzadeh S (2012) Noise control by roughness-induced ground effects and vegetative cover. In: Environmental noise propagation 2012: definitions, measuring and control aspects. Institute of Acoustics, London, pp 16–26

Au WWL (1993) The sonar of dolphins. Springer, New York

Au WWL, Banks K (1998) The acoustics of the snapping shrimp *Synalpheus parneomeris* in Kaneohe Bay. J Acoust Soc Am 103:41–47. https://doi.org/10.1121/1.423234

Aylor D (1972) Sound transmission through vegetation in relation to leaf area density, leaf width, and breadth of canopy. J Acoust Soc Am 51:411–413. https://doi.org/10.1121/1.1912852

Azar JF, Bell D (2016) Acoustic features within a forest community of native and introduced species in New Zealand. Emu 116:22–31. https://doi.org/10.1071/MU14095

Badino A, Borelli D, Gaggero T et al (2012) Noise emitted from ships: impact inside and outside the vessels. Procedia Soc Behav Sci 48:868–879. https://doi.org/10.1016/j.sbspro.2012.06.1064

Bagočius D, Narščius A (2018) Method for the simplistic modelling of the acoustic footprint of the vessels in the shallow marine area. MethodsX 5:1010–1016. https://doi.org/10.1016/j.mex.2018.08.011

Bailey H, Senior B, Simmons D et al (2010) Assessing underwater noise levels during pile-driving at an offshore windfarm and its potential effects on marine mammals. Mar Pollut Bull 60:888–897. https://doi.org/10.1016/j.marpolbul.2010.01.003

Baker MC (2009) Information content in chorus songs of the group-living Australian magpie (*Cracticus tibicen dorsalis*) in Western Australia. Ethology 115:227–238. https://doi.org/10.1111/j.1439-0310.2008.01606.x

Barber JR, Crooks KR, Fristrup KM (2010) The costs of chronic noise exposure for terrestrial organisms. Trends Ecol Evol 25:180–189. https://doi.org/10.1016/j.tree.2009.08.002

Barthel S, Belton S, Raymond CM, Giusti M (2018) Fostering children's connection to nature through authentic situations: the case of saving salamanders at school. Front Psychol 9:928. https://doi.org/10.3389/fpsyg.2018.00928

Basner M, Clark C, Hansell A et al (2017) Aviation noise impacts: state of the science. Noise Health 19:41–50. https://doi.org/10.4103/nah.NAH_104_16

Behr O, van Helversen O (2004) Bat serenades - complex courtship songs of the sac-winged bat (*Saccopteryx bilineata*). Behav Ecol Sociobiol 56:106–115. https://doi.org/10.1007/s00265-004-0768-7

Bennet-Clark HC (1970) The mechanism and efficiency of sound production in mole crickets. J Exp Biol 52:619–652. https://doi.org/10.1242/jeb.52.3.619

Berglund B, Hassmén P, Job RFS (1996) Sources and effects of low-frequency noise. J Acoust Soc Am 99:2985–3002. https://doi.org/10.1121/1.414863

Bermúdez-Cuamatzin E, Delamore Z, Verbeek L et al (2020) Variation in diurnal patterns of singing activity between urban and rural great tits. Front Ecol Evol 8:24. https://doi.org/10.3389/fevo.2020.00246

Bernardini M, Fredianelli L, Fidecaro F et al (2019) Noise assessment of small vessels for action planning in canal cities. Environments 6:31. https://doi.org/10.3390/environments6030031

Bertucci F, Guerra AS, Sturny V et al (2020) A preliminary acoustic evaluation of three sites in the lagoon of Bora Bora, French Polynesia. Environ Biol Fishes 103:891–902. https://doi.org/10.1007/s10641-020-01000-8

Boatman ND, Brickle NW, Hart JD et al (2004) Evidence for the indirect effects of pesticides on farmland birds. Ibis (Lond 1859) 146:131–143. https://doi.org/10.1111/j.1474-919X.2004.00347.x

Boelman NT, Asner GP, Hart PJ, Martin RE (2007) Multitrophic invasion resistance in Hawaii: bioacoustics, field surveys, and airborne remote sensing. Ecol Appl 17:2137–2144. https://doi.org/10.1890/07-0004.1

Boersma HF (1997) Characterization of the natural ambient sound environment: measurements in open agricultural grassland. J Acoust Soc Am 101:2104–2110. https://doi.org/10.1121/1.418141

Bohnenstiehl DR, Lillis A, Eggleston DB (2016) The curious acoustic behavior of estuarine snapping shrimp: temporal patterns of snapping shrimp sound in sub-tidal oyster reef habitat. PLoS One 11:e0143691. https://doi.org/10.1371/journal.pone.0143691

Bolgan M, O'Brien J, Winfield IJ, Gammell M (2016) An investigation of inland water soundscapes: which sonic

sources influence acoustic levels? Proc Meet Acoust 27:070004. https://doi.org/10.1121/2.0000260
Bolgan M, Amorim MCP, Fonseca PJ et al (2018a) Acoustic complexity of vocal fish communities: a field and controlled validation. Sci Rep 8:10559. https://doi.org/10.1038/s41598-018-28771-6
Bolgan M, O'Brien J, Chorazyczewska E et al (2018b) The soundscape of the Arctic Charr spawning grounds in lotic and lentic environments: can passive acoustic monitoring be used to detect spawning activities? Bioacoustics 27:57–85. https://doi.org/10.1080/09524622.2017.1286262
Bolin K (2009) Prediction method for wind-induced vegetation noise. Acta Acust United Acust 95:607–619. https://doi.org/10.3813/AAA.918189
Bond AB, Diamond J (2005) Geographic and ontogenetic variation in the contact calls of the kea (*Nestor notabilis*). Behaviour 142:1–20. https://doi.org/10.1163/1568539053627721
Borelli D, Gaggero T, Rizzuto E, Schenone C (2016) Holistic control of ship noise emissions. Noise Mapp 3:107–119. https://doi.org/10.1515/noise-2016-0008
Bourgeois K, Curé C, Legrand J et al (2007) Morphological versus acoustic analysis: what is the most efficient method for sexing yelkouan shearwaters *Puffinus yelkouan*? J Ornithol 148:261–269. https://doi.org/10.1007/s10336-007-0127-3
Bowling DL, Garcia M, Dunn JC et al (2017) Body size and vocalization in primates and carnivores. Sci Rep 7:41070. https://doi.org/10.1038/srep41070
Bozkurt TS, Demirkale SY (2017) The field study and numerical simulation of industrial noise mapping. J Build Eng 9:60–75. https://doi.org/10.1016/j.jobe.2016.11.007
Brady J (1974) The physiology of insect circadian rhythms. Adv Insect Phys 10:1–115. https://doi.org/10.1016/S0065-2806(08)60129-0
Bregman AS (1990) Auditory scene analysis: the perceptual organization of sound. The MIT Press, Cambridge
Brown A, Garg S, Montgomery J (2019) Automatic rain and cicada chorus filtering of bird acoustic data. Appl Soft Comput 81:105501. https://doi.org/10.1016/j.asoc.2019.105501
Brunetti AE, Saravia AM, Barrionuevo JS, Reichle S (2017) Silent sounds in the Andes: underwater vocalizations of three frog species with reduced tympanic middle ears (Anura: Telmatobiidae: *Telmatobius*). Can J Zool 95:335–343. https://doi.org/10.1139/cjz-2016-0177
Burbidge T, Parson T, Caycedo-Rosales PC et al (2015) Playbacks revisited: asymmetry in behavioural response across an acoustic boundary between two parapatric bird species. Behaviour 152:1933–1951. https://doi.org/10.1163/1568539X-00003309
Buscaino G, Ceraulo M, Pieretti N et al (2016) Temporal patterns in the soundscape of the shallow waters of a Mediterranean marine protected area. Sci Rep 6:34230. https://doi.org/10.1038/srep34230
Buzzetti F, Brizio C, Pavan G (2020) Beyond the audible: wide band (0-125 kHz) field investigation on Italian Orthoptera (Insecta) songs. Biodivers J 11:443–496. https://doi.org/10.31396/Biodiv.Jour.2020.11.2.443.496
Cai M, Yao Y, Wang H (2018) Urban traffic noise maps under 3D complex building environments on a supercomputer. J Adv Transp 2018:7031418. https://doi.org/10.1155/2018/7031418
Caorsi VZ, Both C, Cechin S et al (2017) Effects of traffic noise on the calling behavior of two Neotropical hylid frogs. PLoS One 12:e0183342. https://doi.org/10.1371/journal.pone.0183342
Carson R (1962) Silent Spring. Houghton Mifflin Company, Boston
Caruso F, Alonge G, Bellia G et al (2017) Long-term monitoring of dolphin biosonar activity in deep pelagic waters of the Mediterranean Sea. Sci Rep 7:4321. https://doi.org/10.1038/s41598-017-04608-6
Catchpole CK, Slater PJR (2008) Bird song: biological themes and variations. Cambridge University Press, Cambridge
Cato DH (2008) Ocean ambient noise: its measurement and its significance to marine animals. In: Proceedings of the Institute of Acoustics. Institute of Acoustics, Southampton, pp 1–9
Cerchio S, Dahlheim M (2001) Variation in feeding vocalizations of humpback whales *Megaptera novaeangliae* from southeast Alaska. Bioacoustics 11:277–295. https://doi.org/10.1080/09524622.2001.9753468
Chabert T, Colin A, Aubin T et al (2015) Size does matter: crocodile mothers react more to the voice of smaller offspring. Sci Rep 5:15547. https://doi.org/10.1038/srep15547
Challender DWS, MacMillan DC (2014) Poaching is more than an enforcement problem. Conserv Lett 7:484–494. https://doi.org/10.1111/conl.12082
Chapman NR, Price A (2011) Low frequency deep ocean ambient noise trend in the northeast Pacific Ocean. J Acoust Soc Am 129:EL161–EL165. https://doi.org/10.1121/1.3567084
Charrier I, Mathevon N, Jouventin P, Aubin T (2001) Acoustic communication in a black-headed gull colony: how do chicks identify their parents? Ethology 107:961–974. https://doi.org/10.1046/j.1439-0310.2001.00748.x
Chaudhary A, Burivalova Z, Koh LP, Hellweg S (2016) Impact of forest management on species richness: global meta-analysis and economic trade-offs. Sci Rep 6:23954. https://doi.org/10.1038/srep23954
Cheney DL, Seyfarth RM (1996) Function and intention in the calls of non-human primates. In: Runciman WG, Smith JM, Dunbar RIM (eds) Proceedings of the British Academy, Evolution of social behaviour patterns in primates and man, vol 88. Oxford University Press, Oxford, pp 59–76
Cheney DL, Seyfarth RM (2018) Flexible usage and social function in primate vocalizations. Proc Natl Acad Sci

USA 115:1974–1979. https://doi.org/10.1073/pnas.1717572115
Ciceran M, Murray AM, Rowell G (1994) Natural variation in the temporal patterning of calling song structure in the field cricket *Gryllus pennsylvanicus*: effects of temperature, age, mass, time of day, and nearest neighbor. Can J Zool 72:38–42. https://doi.org/10.1139/z94-006
Clark CW (1982) The acoustic repertoire of the southern right whale, a quantitative analysis. Anim Behav 30: 1060–1071. https://doi.org/10.1016/S0003-3472(82)80196-6
Clark CJ (2021) Ways that animal wings produce sound. Integr Comp Biol 61:696–709. https://doi.org/10.1093/icb/icab008
Clausen KT, Wahlberg M, Beedholm K et al (2010) Click communication in harbour porpoises *Phocoena phocoena*. Bioacoustics 20:1–28. https://doi.org/10.1080/09524622.2011.9753630
Coquereau L, Grall J, Chauvaud L et al (2016) Sound production and associated behaviours of benthic invertebrates from a coastal habitat in the north-east Atlantic. Mar Biol 163:127. https://doi.org/10.1007/s00227-016-2902-2
Courts R, Erbe C, Wellard R et al (2020) Australian long-finned pilot whales (*Globicephala melas*) emit stereotypical, variable, biphonic, multi-component, and sequenced vocalisations, similar to those recorded in the northern hemisphere. Sci Rep 10: 20609. https://doi.org/10.1038/s41598-020-74111-y
Crowley SR, Pietruszka RD (1983) Aggressiveness and vocalization in the leopard lizard (*Gambelia wislizennii*): the influence of temperature. Anim Behav 31:1055–1060. https://doi.org/10.1016/S0003-3472(83)80012-8
Cunnington GM, Fahrig L (2010) Plasticity in the vocalizations of anurans in response to traffic noise. Acta Oecologica 36:463–470. https://doi.org/10.1016/j.actao.2010.06.002
Cure C, Aubin T, Mathevon N (2009) Acoustic convergence and divergence in two sympatric burrowing nocturnal seabirds. Biol J Linn Soc 96:115–134. https://doi.org/10.1111/j.1095-8312.2008.01104.x
Dahl PH, Dall'Osto DR, Farrell DM (2015) The underwater sound field from vibratory pile driving. J Acoust Soc Am 137:3544–3554. https://doi.org/10.1121/1.4921288
Danielsen F, Heegaard M (1995) Impact of logging and plantation development on species diversity: a case study from Sumatra. In: Sandbukt Ø (ed) Management of tropical forests: towards an integrated perspective. Centre for Development and the Environment, University of Oslo, Oslo, pp 73–92
Davies WJ, Adams MD, Bruce NS et al (2013) Perception of soundscapes: an interdisciplinary approach. Appl Acoust 74:224–231. https://doi.org/10.1016/j.apacoust.2012.05.010
De Queiroz MB (2018) How does the zoo soundscape affect the zoo experience for animals and visitors? University of Salford, Manchester
Deichmann JL, Hernández-Serna A, Delagado CJA et al (2017) Soundscape analysis and acoustic monitoring document impacts of natural gas exploration on biodiversity in a tropical forest. Ecol Indic 74:39–48. https://doi.org/10.1016/j.ecolind.2016.11.002
Dentressangle F, Aubin T, Mathevon N (2012) Males use time whereas females prefer harmony: individual call recognition in the dimorphic blue-footed booby. Anim Behav 84:413–420. https://doi.org/10.1016/j.anbehav.2012.05.012
Dey M, Krishnaswamy J, Morisaka T, Kelkar N (2019) Interacting effects of vessel noise and shallow river depth elevate metabolic stress in Ganges river dolphins. Sci Rep 9:15426. https://doi.org/10.1038/s41598-019-51664-1
Di GQ, Lin QL, Li ZG, Kang J (2014) Annoyance and activity disturbance induced by high-speed railway and conventional railway noise: a contrastive case study. Environ Health 13:12. https://doi.org/10.1186/1476-069X-13-12
Dingle C, Halfwerk W, Slabbekoorn H (2008) Habitat-dependent song divergence at subspecies level in the grey-breasted wood-wren. J Evol Biol 21:1079–1089. https://doi.org/10.1111/j.1420-9101.2008.01536.x
Dingle C, Poelstra JW, Halfwerk W et al (2010) Asymmetric response patterns to subspecies-specific song differences in allopatry and parapatry in the gray-breasted wood-wren. Evolution 64:3537–3548. https://doi.org/10.1111/j.1558-5646.2010.01089.x
Doohan B, Fuller S, Parsons S, Peterson EE (2019) The sound of management: acoustic monitoring for agricultural industries. Ecol Indic 96:739–746. https://doi.org/10.1016/j.ecolind.2018.09.029
Dragoset B (2000) Introduction to air guns and air-gun arrays. Lead Edge 19:892–897. https://doi.org/10.1190/1.1438741
Drozdova L, Butorina M, Kuklin D (2019) Evaluation and reduction of the common effect of road and rail noise. In: Proceedings of the 26th International Congress on Sound and Vibration. International Institute of Acoustics and Vibration, Montreal
Duarte MHL, Caliari EP, Scarpelli MDA et al (2019) Effects of mining truck traffic on cricket calling activity. J Acoust Soc Am 146:656–664. https://doi.org/10.1121/1.5119125
Duffy JE (1996) Eusociality in coral-reef shrimp. Nature 381:512–514. https://doi.org/10.1038/381512a0
Duffy JE, Macdonald KS (1999) Colony structure of the social snapping shrimp *Synalpheus filidigitus* in Belize. J Crustac Biol 19:283–292. https://doi.org/10.1163/193724099X00097
Dulvy NK, Fowler SL, Musick JA et al (2014) Extinction risk and conservation of the world's sharks and rays. Elife 3:e00590. https://doi.org/10.7554/eLife.00590
Duncan AJ, Maggi AL (2006) A consistent, user friendly interface for running a variety of underwater acoustic

propagation codes. In: Proceedings of Acoustics. Christchurch, 20–22 November 2006

Dunlop RA, Noad MJ, Cato DH, Stokes D (2007) The social vocalization repertoire of east Australian migrating humpback whales (*Megaptera novaeangliae*). J Acoust Soc Am 122:2893–2905. https://doi.org/10.1121/1.2783115

Duque FG, Rodríguez-Saltos CA, Wilczynski W (2018) High-frequency vocalizations in Andean hummingbirds. Curr Biol 28:927–928. https://doi.org/10.1016/j.cub.2018.07.058

Dziak RP, Fox CG (2002) Evidence of harmonic tremor from a submarine volcano detected across the Pacific Ocean basin. J Geophys Res 107:1–11. https://doi.org/10.1029/2001JB000177

Dziak RP, Haxel JH, Matsumoto H et al (2017) Ambient sound at challenger deep, mariana trench. Oceanography 30:186–197. https://doi.org/10.5670/oceanog.2017.240

Erbe C (2009) Underwater noise from pile driving in Moreton Bay, Qld. Acoust Aust 37:87–92

Erbe C (2013) Underwater noise of small personal watercraft (jet skis). J Acoust Soc Am 133:EL326–EL330. https://doi.org/10.1121/1.4795220

Erbe C, King AR (2009) Modelling cumulative sound exposure around marine seismic surveys. J Acoust Soc Am 125:2443–2451. https://doi.org/10.1121/1.3089588

Erbe C, McPherson C (2012) Acoustic characterisation of bycatch mitigation pingers on shark control nets in Queensland, Australia. Endanger Species Res 19:109–121. https://doi.org/10.3354/esr00467

Erbe C, McPherson C (2017) Underwater noise from geotechnical drilling and standard penetration testing. J Acoust Soc Am 142:EL281–EL285. https://doi.org/10.1121/1.5003328

Erbe C, MacGillivray A, Williams R (2012) Mapping cumulative noise from shipping to inform marine spatial planning. J Acoust Soc Am 132:EL423–EL428. https://doi.org/10.1121/1.4758779

Erbe C, McCauley RD, McPherson C, Gavrilov A (2013) Underwater noise from offshore oil production vessels. J Acoust Soc Am 133:EL465–EL470. https://doi.org/10.1121/1.4802183

Erbe C, Williams R, Sandilands D, Ashe E (2014) Identifying modeled ship noise hotspots for marine mammals of Canada's Pacific region. PLoS One 9:e89820. https://doi.org/10.1371/journal.pone.0089820

Erbe C, Verma A, McCauley R et al (2015) The marine soundscape of the Perth Canyon. Prog Oceanogr 137:38–51. https://doi.org/10.1016/j.pocean.2015.05.015

Erbe C, Liong S, Koessler MW et al (2016a) Underwater sound of rigid-hulled inflatable boats. J Acoust Soc Am 139:EL223–EL227. https://doi.org/10.1121/1.4954411

Erbe C, McCauley R, Gavrilov A et al (2016b) The underwater soundscape around Australia. In: Proceedings of Acoustics. Brisbane, 9–11 November 2016.

Erbe C, Parsons M, Duncan AJ, Allen K (2016c) Underwater acoustic signatures of recreational swimmers, divers, surfers and kayakers. Acoust Aust 44:333–341. https://doi.org/10.1007/s40857-016-0062-7

Erbe C, Wintner S, Dudley SFJ, Plön S (2016d) Revisiting acoustic deterrence devices: long-term bycatch data from South Africa's bather protection nets. Proc Meet Acoust 27:010025. https://doi.org/10.1121/2.0000306

Erbe C, Dunlop R, Jenner KCS et al (2017) Review of underwater and in-air sounds emitted by Australian and Antarctic marine mammals. Acoust Aust 45:179–241. https://doi.org/10.1007/s40857-017-0101-z

Erbe C, Williams R, Parsons M et al (2018) Underwater noise from airplanes: an overlooked source of ocean noise. Mar Pollut Bull 137:656–661. https://doi.org/10.1016/j.marpolbul.2018.10.064

Erbe C, Marley SA, Schoeman RP et al (2019) The effects of ship noise on marine mammals: a review. Front Mar Sci 6:606. https://doi.org/10.3389/fmars.2019.00606

Erbe C, Schoeman RP, Peel D, Smith JN (2021) It often howls more than it chugs: wind versus ship noise under water in Australia's maritime regions. J Mar Sci Eng 9:1–27. https://doi.org/10.3390/jmse9050472

Erisman BE, Rowell TJ (2017) A sound worth saving: acoustic characteristics of a massive fish spawning aggregation. Biol Lett 13:20170656. https://doi.org/10.1098/rsbl.2017.0656

Ernstes R, Quinn JE (2016) Variation in bird vocalizations across a gradient of traffic noise as a measure of an altered urban soundscape. Cities Environ 8:7

European Environment Agency [EEA] (2014) Noise in Europe. Publications Office of the European Union, Luxembourg

Farina A, Gage SH (2017) Ecoacoustics: the ecological role of sounds. Wiley, Hoboken

Farina A, Righini R, Fuller S et al (2021) Acoustic complexity indices reveal the acoustic communities of the old-growth Mediterranean forest of Sasso Fratino Integral Natural Reserve (Central Italy). Ecol Indic 120:106927. https://doi.org/10.1016/j.ecolind.2020.106927

Feng AS, Narins PM, Xu CH et al (2006) Ultrasonic communication in frogs. Nature 440:333–336. https://doi.org/10.1038/nature04416

Fenton MB, Portfors CV, Rautenback IL, Waterman JM (1998) Compromises: sound frequencies used in echolocation by aerial-feeding bats. Can J Zool 76:1174–1182. https://doi.org/10.1139/z98-043

Fernández-Juricic E, Campagna C, Enriquez V, Ortiz CL (1999) Vocal communication and individual variation in breeding south American sea lions. Behaviour 136:495–517. https://doi.org/10.1163/156853999501441

Ferreira LM, Oliveira EG, Lopes LC et al (2018) What do insects, anurans, birds, and mammals have to say about soundscape indices in a tropical savanna. J Ecoacoust 2: #PVH6YZ. https://doi.org/10.22261/jea.pvh6yz

Fichtel C (2020) Monkey alarm calling: it ain't all referential, or is it? Anim Behav Cogn 7:101–107. https://doi.org/10.26451/abc.07.02.04.2020

Findlay CR, Ripple HD, Coomber F et al (2018) Mapping widespread and increasing underwater noise pollution from acoustic deterrent devices. Mar Pollut Bull 135: 1042–1050. https://doi.org/10.1016/j.marpolbul.2018.08.042

Flint EL, Minot EO, Perry PE, Stafford KJ (2014) A survey of public attitudes towards barking dogs in New Zealand. N Z Vet J 62:321–327. https://doi.org/10.1080/00480169.2014.921852

Fournet MEH, Gabriele CM, Sharpe F et al (2018) Feeding calls produced by solitary humpback whales. Mar Mamm Sci 34:851–865. https://doi.org/10.1111/mms.12485

Fox CG, Matsumoto H, Lau TKA (2001) Monitoring Pacific Ocean seismicity from an autonomous hydrophone array. J Geophys Res 106:4183–4206. https://doi.org/10.1029/2000JB900404

Francis CD, Newman P, Taff BD et al (2017) Acoustic environments matter: synergistic benefits to humans and ecological communities. J Environ Manag 203: 245–254. https://doi.org/10.1016/j.jenvman.2017.07.041

Franco LS, Shanahan DF, Fuller RA (2017) A review of the benefits of nature experiences: more than meets the eye. Int J Environ Res Public Health 14:864. https://doi.org/10.3390/ijerph14080864

Freeman SE, Freeman LA, Giorli G, Haas AF (2018) Photosynthesis by marine algae produces sound, contributing to the daytime soundscape on coral reefs. PLoS One 13:e0201766. https://doi.org/10.1371/journal.pone.0201766

Gadziola MA, Grimsley JMS, Faure PA, Wenstrup JJ (2012) Social vocalizations of big brown bats vary with behavioral context. PLoS One 7:e44550. https://doi.org/10.1371/journal.pone.0044550

Gage SH, Axel AC (2014) Visualization of temporal change in soundscape power of a Michigan lake habitat over a 4-year period. Ecol Inform 21:100–109. https://doi.org/10.1016/j.ecoinf.2013.11.004

Galeotti P, Sacchi R, Fasola M, Ballasina D (2005) Do mounting vocalisations in tortoises have a communication function? A comparative analysis. Herpetol J 15:61–71

Garrett JK, Blondel P, Godley BJ et al (2016) Long-term underwater sound measurements in the shipping noise indicator bands 63 Hz and 125 Hz from the port of Falmouth Bay, UK. Mar Pollut Bull 110:438–448. https://doi.org/10.1016/j.marpolbul.2016.06.021

Gavrilov A (2018) Propagation of underwater noise from an offshore seismic survey in Australia to Antarctica: measurements and modelling. Acoust Aust 46:143–149. https://doi.org/10.1007/s40857-018-0131-1

Gavrilov AN, Parsons MJG (2014) A MATLAB tool for the characterisation of recorded underwater sounds (CHORUS). Acoust Aust 42:190–196

Gavrilov AN, McCauley RD, Gedamke J (2012) Steady inter and intra-annual decrease in the vocalization frequency of Antarctic blue whales. J Acoust Soc Am 131:4476–4480. https://doi.org/10.1121/1.4707425

Gazioğlu C, Müftüoğlu AE, Demir V et al (2015) Connection between ocean acidification and sound propagation. Int J Environ Geoinformatics 2:16–26. https://doi.org/10.30897/ijegeo.303538

Gerstein ER, Trygonis V, McCulloch S et al (2014) Female north Atlantic right whales produce gunshot sounds. J Acoust Soc Am 135:2369. https://doi.org/10.1121/1.4877814

Gibbs JP, Breisch AR (2001) Climate warming and calling phenology of frogs near Ithaca, New York, 1900-1999. Conserv Biol 15:1175–1178. https://doi.org/10.1046/j.1523-1739.2001.0150041175.x

Gil D, Honarmand M, Pascual J et al (2015) Birds living near airports advance their dawn chorus and reduce overlap with aircraft noise. Behav Ecol 26:435–443. https://doi.org/10.1093/beheco/aru207

Giles JC, Davis JA, McCauley RD, Kuchling G (2009) Voice of the turtle: the underwater acoustic repertoire of the long-necked freshwater turtle, *Chelodina oblonga*. J Acoust Soc Am 126:434–443. https://doi.org/10.1121/1.3148209

Gill SA, Bierema AMK (2013) On the meaning of alarm calls: a review of functional reference in avian alarm calling. Ethology 119:449–461. https://doi.org/10.1111/eth.12097

Gordon J, Gillespie D, Potter J et al (2003) A review of the effects of seismic surveys on marine mammals. Mar Technol Soc J 37:16–34. https://doi.org/10.4031/002533203787536998

Gordon TAC, Radford AN, Davidson IK et al (2019) Acoustic enrichment can enhance fish community development on degraded coral reef habitat. Nat Commun 10:5414. https://doi.org/10.1038/s41467-019-13186-2

Gottesman BL, Francomano D, Zhao Z et al (2020) Acoustic monitoring reveals diversity and surprising dynamics in tropical freshwater soundscapes. Freshw Biol 65:117–132. https://doi.org/10.1111/fwb.13096

Grafe TU (2005) Anuran choruses as communication networks. In: McGregor G (ed) Animal communication networks. Cambridge University Press, Cambridge, pp 277–299

Greta M, Louena S, Arianna A, et al (2019) Prediction of off-site noise levels reduction in open-air music events within densely populated urban areas. In: INTER-NOISE 2019 MADRID - 48th International Congress and Exhibition on Noise Control Engineering. International Institute of Noise Control Engineering, Madrid, 16–19 June 2019

Gulyas K, Pinte G, Augusztinovicz F, et al (2002) Active noise control in agricultural machines. In: Proceedings of the 2002 International Conference on Noise and Vibration Engineering, ISMA, Leuven, 16–18 September 2002

Halfwerk W, Slabbekoorn H (2009) A behavioural mechanism explaining noise-dependent frequency use in urban birdsong. Anim Behav 78:1301–1307. https://doi.org/10.1016/j.anbehav.2009.09.015

Harris SA, Shears NT, Radford CA (2016) Ecoacoustic indices as proxies for biodiversity on temperate reefs. Methods Ecol Evol 7:713–724. https://doi.org/10.1111/2041-210X.12527

Hart PJ, Hall R, Ray W et al (2015) Cicadas impact bird communication in a noisy tropical rainforest. Behav Ecol 26:839–842. https://doi.org/10.1093/beheco/arv018

Hatch L, Clark C, Merrick R et al (2008) Characterizing the relative contributions of large vessels to total ocean noise fields: a case study using the Gerry E. Studds Stellwagen Bank National Marine Sanctuary. Environ Manag 42:735–752. https://doi.org/10.1007/s00267-008-9169-4

Hauser DDW, Laidre KL, Stafford KM et al (2016) Decadal shifts in autumn migration timing by Pacific Arctic beluga whales are related to delayed annual sea ice formation. Glob Change Biol 23:2206–2217. https://doi.org/10.1111/gcb.13564

Haver SM, Klinck H, Nieukirk SL et al (2017) The not-so-silent world: measuring Arctic, Equatorial and Antarctic soundscapes in the Atlantic Ocean. Deep Res I Oceanogr Res Pap 122:95–104. https://doi.org/10.1016/j.dsr.2017.03.002

Herberholz J, Schmitz B (1999) Flow visualisation and high speed video analysis of water jets in snapping shrimp (*Alpheus heterochaelis*). J Comp Physiol A 185:41–49. https://doi.org/10.1007/s003590050364

Herman LM (2017) The multiple functions of male song within the humpback whale (*Megaptera novaeangliae*) mating system: review, evaluation, and synthesis. Biol Rev 92:1795–1818. https://doi.org/10.1111/brv.12309

Hermannsen L, Beedholm K, Tougaard J, Madsen PT (2014) High frequency components of ship noise in shallow water with a discussion of implications for harbor porpoises (*Phocoena phocoena*). J Acoust Soc Am 136:1640–1653. https://doi.org/10.1121/1.4893908

Hermannsen L, Tougaard J, Beedholm K et al (2015) Characteristics and propagation of airgun pulses in shallow water with implications for effects on small marine mammals. PLoS One 10:e0133436. https://doi.org/10.1371/journal.pone.0133436

Herzing DL (1996) Vocalizations and associated underwater behavior of free-ranging Atlantic spotted dolphins, *Stenella frontalis* and bottlenose dolphins, *Tursiops truncatus*. Aquat Mamm 22:61–79. https://doi.org/10.12966/abc.02.02.2015

Hiley HM, Perry S, Hartley S, King SL (2017) What's occurring? Ultrasonic signature whistle use in Welsh bottlenose dolphins (*Tursiops truncatus*). Bioacoustics 26:25–35. https://doi.org/10.1080/09524622.2016.1174885

Hoffmann M, Belant JL, Chanson JS et al (2011) The changing fates of the world's mammals. Philos Trans R Soc B 366:2598–2610. https://doi.org/10.1098/rstb.2011.0116

Hole DG, Perkins AJ, Wilson JD et al (2005) Does organic farming benefit biodiversity? Biol Conserv 122:113–130. https://doi.org/10.1016/j.biocon.2004.07.018

Holt DE, Johnston CE (2015) Traffic noise masks acoustic signals of freshwater stream fish. Biol Conserv 187:27–33. https://doi.org/10.1016/j.biocon.2015.04.004

Insley SJ, Phillips AV, Charrier I (2010) A review of social recognition in pinnipeds. Aquat Mamm 29:181–201. https://doi.org/10.1578/016754203101024149

Intergovernmental Panel on Climate Change [IPCC] (2014) Climate change 2014 synthesis report. Contribution of working groups I, II and III on the fifth assessment report of the intergovernmental panel on climate change. IPCC, Geneva

Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services [IPBES] (2019) Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the intergovernmental science-policy platform on biodiversity and ecosystem services. PBES Secretariat, Bonn

International Organization for Standardization [ISO] (2014) International Standard 12913-1 acoustics - soundscape - Part 1: definition and conceptual framework. International Organization for Standardization, Geneva

International Organization for Standardization [ISO] (2017) International Standard 18405 underwater acoustics - terminology. International Organization for Standardization, Geneva

Iversen RTS, Perkins PJ, Dionne RD (1963) An indication of underwater sound production by squid. Nature 199:250–251. https://doi.org/10.1038/199250a0

Jacobs SR, Terhune JM (2002) The effectiveness of acoustic harassment devices in the Bay of Fundy, Canada: seal reactions and a noise exposure model. Aquat Mamm 28:147–158

Jahn O, Ganchev TD, Marques MI, Schuchmann KL (2017) Automated sound recognition provides insights into the behavioral ecology of a tropical bird. PLoS One 12:e0169041. https://doi.org/10.1371/journal.pone.0169041

Jézéquel Y, Bonnel J, Coston-Guarini J, Chauvaud L (2019) Revisiting the bioacoustics of European spiny lobsters *Palinurus elephas*: comparison of antennal rasps in tanks and *in situ*. Mar Ecol Prog Ser 615:143–157. https://doi.org/10.3354/meps12935

Johnson JB, Lees JM, Yepes H (2006) Volcanic eruptions, lightning, and a waterfall: differentiating the menagerie of infrasound in the Ecuadorian jungle. Geophys Res Lett 33:L06308. https://doi.org/10.1029/2005GL025515

Joo W, Gage SH, Kasten EP (2011) Analysis and interpretation of variability in soundscapes along an urban–rural gradient. Landsc Urban Plan 103:259–276. https://doi.org/10.1016/j.landurbplan.2011.08.001

Kaatz IM (2002) Multiple sound-producing mechanisms in teleost fishes and hypotheses regarding their

behavioural significance. Bioacoustics 12:230–233. https://doi.org/10.1080/09524622.2002.9753705
Kaatz IM (2011) How fishes use sound: quiet to loud and simple to complex signalling. In: Farrell AP (ed) Encyclopedia of fish physiology. Elsevier, San Diego, pp 684–691
Kaiser F, Rohde T (2013) Orlando theme park acoustics - a soundscape analysis. In: Internoise. International Congress and Exposition on Noise Control Engineering, 15–18 September 2013, Austrian Noise Abatement Association, Innsbruck
Kariel HG (1990) Factors affecting response to noise in outdoor recreational environments. Can Geogr 34: 142–149. https://doi.org/10.1111/j.1541-0064.1990.tb01259.x
Kasten EP, Gage SH, Fox J, Joo W (2012) The remote environmental assessment laboratory's acoustic library: an archive for studying soundscape ecology. Ecol Inform 12:50–67. https://doi.org/10.1016/j.ecoinf.2012.08.001
Kasumyan AO (2008) Sounds and sound production in fishes. J Ichthyol 48:981–1030. https://doi.org/10.1134/S0032945208110039
Katz J, Hafner SD, Donovan T (2016) Tools for automated acoustic monitorig with the R package monitoR. Bioacoustics 25:197–210. https://doi.org/10.1080/09524622.2016.1138415
Kawakita S, Ichikawa K (2019) Automated classification of bees and hornet using acoustic analysis of their flight sounds. Apidologie 50:71–79. https://doi.org/10.1007/s13592-018-0619-6
Kerr JT, Cihlar J (2004) Patterns and causes of species endangerment in Canada. Ecol Appl 14:743–753. https://doi.org/10.1890/02-5117
Kleijn D, Kohler F, Báldi A et al (2009) On the relationship between farmland biodiversity and land-use intensity in Europe. Proc R Soc B 276:903–909. https://doi.org/10.1098/rspb.2008.1509
Kleijn D, Rundlöf M, Scheper J et al (2011) Does conservation on farmland contribute to halting the biodiversity decline? Trends Ecol Evol 26:474–481. https://doi.org/10.1016/j.tree.2011.05.009
Klenova AV (2015) Chick begging calls reflect degree of hunger in three Auk species (Charadriiformes: Alcidae). PLoS One 10:e0140151. https://doi.org/10.1371/journal.pone.0140151
Klinck H, Nieukirk SL, Mellinger DK et al (2012) Seasonal presence of cetaceans and ambient noise levels in polar waters of the north Atlantic. J Acoust Soc Am 132:EL176–EL181. https://doi.org/10.1121/1.4740226
Kline LR, DeAngelis AI, McBride C et al (2020) Sleuthing with sound: understanding vessel activity in marine protected areas using passive acoustic monitoring. Mar Policy 120:104138. https://doi.org/10.1016/j.marpol.2020.104138
Kloepper LN, Simmons AM (2014) Bioacoustic monitoring contributes to an understanding of climate change. Acoust Tod 10:8–15
Knight L, Ladich F (2014) Distress sounds of thorny catfishes emitted underwater and in air: characteristics and potential significance. J Exp Biol 217:4068–4078. https://doi.org/10.1242/jeb.110957
Knowlton RE, Moulton JM (1963) Sound production in the snapping shrimps *Alpheus* (*Crangon*) and *Synalpheus*. Biol Bull 125:311–331. https://doi.org/10.2307/1539406
Knudsen VO, Alford RS, Emling JW (1948) Underwater ambient noise. J Mar Res 7:410–429
Koschinski S, Culik BM, Henriksen OD et al (2003) Behavioural reactions of free-ranging propoises and seals to the noise of a simulated 2 MW windpower generator. Mar Ecol Prog Ser 265:263–273. https://doi.org/10.3354/meps265263
Krause BL (1987) Bio-acoustics: habitat ambience & ecological balance. Signal 57:14–16
Krause BL (1993) The niche hypothesis: a virtual symphony of animal sounds, the origins of musical expression and the health of habitats. Soundscape Newsl 6:6–10
Krause B (2008) Anatomy of the soundscape: evolving perspectives. J Audio Eng Soc 56:73–80
Krause B (2012) The great animal orchestra. Little, Brown and Company, Boston
Krause B, Farina A (2016) Using ecoacoustic methods to survey the impacts of climate change on biodiversity. Biol Conserv 195:245–254. https://doi.org/10.1016/j.biocon.2016.01.013
Kruger DJD, Du Preez LH (2016) The effect of airplane noise on frogs: a case study on the critically endangered Pickersgill's reed frog (*Hyperolius pickersgilli*). Ecol Res 31:393–405. https://doi.org/10.1007/s11284-016-1349-8
Kuehne LM, Padgham BL, Olden JD (2013) The soundscapes of lakes across an urbanization gradient. PLoS One 8:e55661. https://doi.org/10.1371/journal.pone.0055661
Kukulski B, Wszolek T, Mleczko D (2018) The impact of fireworks noise on the acoustic climate in urban areas. Arch Acoust 43:697–705. https://doi.org/10.24425/aoa.2018.125163
Ladich F (1997) Agonistic behaviour and significance of sounds in vocalizing fish. Mar Freshw Behav Physiol 29:87–108. https://doi.org/10.1080/10236249709379002
Ladich F, Winkler H (2017) Acoustic communication in terrestrial and aquatic vertebrates. J Exp Biol 220: 2306–2317. https://doi.org/10.1242/jeb.132944
LaManna JA, Martin TE (2017) Logging impacts on avian species richness and composition differ across latitudes and foraging and breeding habitat preferences. Biol Rev 92:1657–1674. https://doi.org/10.1111/brv.12300
Larom D, Garstang M, Payne K et al (1997) The influence of surface atmospheric conditions on the range and area reached by animal vocalizations. J Exp Biol 200: 421–431
Lattenkamp EZ, Shields SM, Schutte M et al (2019) The vocal repertoire of pale spear-nosed bats in a social

roosting context. Front Ecol Evol 7:116. https://doi.org/10.3389/fevo.2019.00116
Lengagne T, Slater PJB (2002) The effects of rain on acoustic communication: Tawny owls have good reason for calling less in wet weather. Proc R Soc London B 269:2121–2125. https://doi.org/10.1098/rspb.2002.2115
Lengagne T, Jouventin P, Aubin T (1999) Finding one's mate in a king penguin colony: efficiency of acoustic communication. Behaviour 136:833–846. https://doi.org/10.1163/156853999501595
Lewicki MS, Olshausen BA, Surlykke A, Moss CF (2014) Scene analysis in the natural environment. Front Psychol 5:199. https://doi.org/10.3389/fpsyg.2014.00199
Lillis A, Perelman JN, Panyi A, Mooney TA (2017) Sound production patterns of big-clawed snapping shrimp (*Alpheus* spp.) are influenced by time-of-day and social context. J Acoust Soc Am 142:3311–3320. https://doi.org/10.1121/1.5012751
Linke S, Gifford T, Desjonquères C (2020) Six steps towards operationalising freshwater ecoacoustic monitoring. Freshw Biol 65:1–6. https://doi.org/10.1111/fwb.13426
Lorang MS, Tonolla D (2014) Combining active and passive hydroacoustic techniques during flood events for rapid spatial mapping of bedload transport patterns in gravel-bed rivers. Fundam Appl Limnol 184:231–246. https://doi.org/10.1127/1863-9135/2014/0552
Ma BB, Nystuen JA, Lien RC (2005) Prediction of underwater sound levels from rain and wind. J Acoust Soc Am 117:3555–3565. https://doi.org/10.1121/1.1910283
MacGillivray A, de Jong C (2021) A reference spectrum model for estimating source levels of marine shipping based on Automated Identification System data. J Mar Sci Eng 9:369. https://doi.org/10.3390/jmse9040369
Macharia JM, Thenya T, Ndiritu GG (2010) Management of highland wetlands in central Kenya: the importance of community education, awareness and eco-tourism in biodiversity conservation. Biodiversity 11:85–90. https://doi.org/10.1080/14888386.2010.9712652
Mack AL, Jones J (2003) Low-frequency vocalizations by cassowaries (*Casuarius* spp.). Auk 120:1062–1068. https://doi.org/10.1093/auk/120.4.1062
Maglio A, Soares C, Bouzidi M et al (2015) Mapping shipping noise in the Pelagos Sanctuary (French part) through acoustic modelling to assess potential impacts on marine mammals. Sci Rep Port Cros Natl Park 29:167–185
Marchal J, Fabianek F, Scott C et al (2020) "bioacoustics": analyse audio recordings and automatically extract animal vocalizations. R package version 0.2.8
Marley SA, Erbe C, Salgado-Kent CP (2016) Underwater sound in an urban estuarine river: sound sources, soundscape contribution, and temporal variability. Acoust Aust 44:171–186. https://doi.org/10.1007/s40857-015-0038-z
Martin SB, Cott PA (2016) The under-ice soundscape in Great Slave Lake near the city of Yellowknife, Northwest Territories, Canada. J Great Lakes Res 42:248–255. https://doi.org/10.1016/j.jglr.2015.09.012
Martin SB, Popper AN (2016) Short- and long-term monitoring of underwater sound levels in the Hudson River (New York, USA). J Acoust Soc Am 139:1886–1897. https://doi.org/10.1121/1.4944876
Mason NA, Burns KJ (2015) The effect of habitat and body size on the evolution of vocal displays in Traupidae (tanagers), the largest family of songbirds. Biol J Linn Soc 114:538–551. https://doi.org/10.1111/bij.12455
Matsumoto H, Bohnenstiel DR, Tournadre J et al (2014) Antarctic icebergs: a significant natural ocean sound source in the southern hemisphere. Geochem Geophys Geosyst 15:3448–3458. https://doi.org/10.1002/2014GC005454
McKenna MF, Ross D, Wiggins SM, Hildebrand JA (2012) Underwater radiated noise from modern commercial ships. J Acoust Soc Am 131:92–103. https://doi.org/10.1121/1.3664100
McShane LJ, Estes JA, Riedman ML, Staedler MM (1995) Repertoire, structure, and individual variation of vocalizations in the sea otter. J Mammol 76:414–427. https://doi.org/10.2307/1382352
McWilliam JN, Hawkins AD (2013) A comparison of inshore marine soundscapes. J Exp Mar Bio Ecol 446:166–176. https://doi.org/10.1016/j.jembe.2013.05.012
McWilliam JN, McCauley RD, Erbe C, Parsons MJG (2017) Patterns of biophonic periodicity on coral reefs in the Great Barrier Reef. Sci Rep 7:17459. https://doi.org/10.1038/s41598-017-15838-z
Mellinger DK, Clark CW (2003) Blue whale (*Balaenoptera musculus*) sounds from the north Atlantic. J Acoust Soc Am 114:1108–1119. https://doi.org/10.1121/1.1593066
Mercado IIIE (2018) The sonar model for humpback whale song revised. Front Psychol 9:1156. https://doi.org/10.3389/fpsyg.2018.01156
Merchant ND, Barton TR, Thompson PM et al (2013) Spectral probability density as a tool for ambient noise analysis. J Acoust Soc Am 133:EL262–EL267. https://doi.org/10.1121/1.4794934
Merchant ND, Fristrup KM, Johnson MP et al (2015) Measuring acoustic habitats. Methods Ecol Evol 6:257–265. https://doi.org/10.1111/2041-210X.12330
Mikhalevsky PN (2001) Acoustics, Arctic. In: Steele JH, Thorpe SA, Turekian KK (eds) Encyclopedia of ocean sciences, vol 1. Elsevier, San Diego, pp 53–61
Miksis-Olds JL, Bradley DL, Maggie Niu X (2013) Decadal trends in Indian Ocean ambient sound. J Acoust Soc Am 134:3464–3475. https://doi.org/10.1121/1.4821537
Mooney TA, Di Iorio L, Lammers M et al (2020) Listening forward: approaching marine biodiversity assessments using acoustic methods. R Soc Open Sci 7:201287. https://doi.org/10.1098/rsos.201287

Morton ES (1975) Ecological sources of selection on avian sounds. Am Nat 109:17–34

Mulard H, Aubin T, White JF et al (2009) Voice variance may signify ongoing divergence among black-legged kittiwake populations. Biol J Linn Soc 97:289–297. https://doi.org/10.1111/j.1095-8312.2009.01198.x

Mullet TC, Gage SH, Morton JM, Huettmann F (2016) Temporal and spatial variation of a winter soundscape in south-central Alaska. Landsc Ecol 31:1117–1137. https://doi.org/10.1007/s10980-015-0323-0

Mullet TC, Farina A, Gage SH (2017a) The acoustic habitat hypothesis: an ecoacoustics perspective on species habitat selection. Biosemiotics 10:319–336. https://doi.org/10.1007/s12304-017-9288-5

Mullet TC, Morton JM, Gage SH, Huettmann F (2017b) Acoustic footprint of snowmobile noise and natural quiet refugia in an Alaskan wilderness. Nat Areas J 37:332–349. https://doi.org/10.3375/043.037.0308

Mumm CAS, Knörnschild M (2014) The vocal repertoire of adult and neonate giant otters (*Pteronura brasiliensis*). PLoS One 9:e112562. https://doi.org/10.1371/journal.pone.0112562

Narins PM (1990) Seismic communication in anuran amphibians. Bioscience 40:268–274. https://doi.org/10.2307/1311263

Narins PM, Meenderink WF (2014) Climate change and frog calls: long-term correlations along a tropical altitudinal gradient. Proc R Soc B 281:20140401. https://doi.org/10.1098/rspb.2014.0401

National Park Service [NPS] (2000) Director's order #47: soundscape preservation and noise management. https://www.nps.gov/policy/DOrders/DOrder47.html. Accessed 5 Feb 2020

Nieukirk SL, Stafford KM, Mellinger DK et al (2004) Low-frequency whale and seismic airgun sounds recorded in the mid-Atlantic Ocean. J Acoust Soc Am 115:1832–1843. https://doi.org/10.1121/1.1675816

Nieukirk SL, Mellinger DK, Moore SE et al (2012) Sounds from airguns and fin whales recorded in the mid-Atlantic Ocean, 1999-2009. J Acoust Soc Am 131:1102–1112. https://doi.org/10.1121/1.3672648

Nijman V (2001) Effect of behavioural changes due to habitat disturbance on density estimation of rain forest vertebrates, as illustrated by gibbons (Primates: Hylobatidae). In: Hillegers PJM, de Iongh HH (eds) The balance between biodiversity conservation and sustainable use of tropical rain forests. Tropenbos Foundation, Wageningen, pp 217–226

Nityananda V, Bee MA (2011) Finding your mate at a cocktail party: frequency separation promotes auditory stream segregation of concurrent voices in multi-species frog choruses. PLoS One 6:e21191. https://doi.org/10.1371/journal.pone.0021191

Nystuen JA (1986) Rainfall measurements using underwater ambient noise. J Acoust Soc Am 79:972–982. https://doi.org/10.1121/1.393695

O'Neal D (1998) Comparison of the underwater ambient noise measured in three large exhibits at the Monterey Bay aquarium and in the inner Monterey Bay. M.Sc. thesis, Naval Postgraduate School, Monterey. Available from https://apps.dtic.mil/sti/pdfs/ADA350428.pdf (accessed on 21 June 2022)

Obrist MK, Pavan G, Sueur J et al (2010) Bioacoustics approaches in biodiversity inventories. In: Eymann J, Degreef J, Häuser C et al (eds) Manual on field recording techniques and protocols for all taxa biodiverity inventories. Abc Taxa, Brussels, pp 68–99

O'Connell-Rodwell CE, Arnason BT, Hart LA (2000) Seismic properties of Asian elephant (*Elephas maximus*) vocalizations and locomotion. J Acoust Soc Am 108:3066–3072. https://doi.org/10.1121/1.1323460

Odom KJ, Hall ML, Riebel K et al (2014) Female song is widespread and ancestral in songbirds. Nat Commun 5:3379. https://doi.org/10.1038/ncomms4379

Owen MA, Swaisgood RR, Czekala NM et al (2004) Monitoring stress in captive giant pandas (*Ailuropoda melanoleuca*): behavioral and hormonal response to ambient noise. Zoo Biol 23:147–164. https://doi.org/10.1002/zoo.10124

Padua SM (1994) Conservationn awareness through an environmental education programme in the Atlantic forest of Brazil. Environ Conserv 21:145–151. https://doi.org/10.1017/S0376892900024577

Pagniello CMLS, Cimino MA, Terrill E (2019) Mapping fish chorus distributions in southern California using an autonomous wave glider. Front Mar Sci 6:526. https://doi.org/10.3389/fmars.2019.00526

Parks SE, Hamilton PK, Kraus SD, Tyack PL (2006) The gunshot sound produced by male north Atlantic right whales (*Eubalaena glacialis*) and its potential function in reproductive advertisement. Mar Mamm Sci 21:458–475. https://doi.org/10.1111/j.1748-7692.2005.tb01244.x

Parks SE, Clark CW, Tyack PL (2007) Short- and long-term changes in right whale calling behavior: the potential effects of noise on acoustic communication. J Acoust Soc Am 122:3725–3731. https://doi.org/10.1121/1.2799904

Parks SE, Miksis-Olds JL, Denes SL (2014) Assessing marine ecosystem acoustic diversity across ocean basins. Ecol Inform 21:81–88. https://doi.org/10.1016/j.ecoinf.2013.11.003

Parsons MJG, Salgado-Kent CP, Marley SA et al (2016) Characterizing diversity and variation in fish choruses in Darwin Harbour. ICES J Mar Sci 73:2058–2074. https://doi.org/10.1093/icesjms/fsw037

Parsons MJG, Duncan AJ, Parsons SK, Erbe C (2020) Reducing vessel noise: an example of a solar-electric passenger ferry. J Acoust Soc Am 147:3575–3583. https://doi.org/10.1121/10.0001264

Parsons MJG, Erbe C, Meekan MG, Parsons SK (2021) A review and meta-analysis of underwater noise radiated by small (<25 m length ) vessels. J Mar Sci Eng 9:827. https://doi.org/10.3390/jmse9080827

Pavan G (2017) Fundamentals of soundscape conservation. In: Farina A, Gage SH (eds)

Ecoacoustics. The ecological role of sound. Wiley, Hoboken, pp 235–258
Pavan G, Priano M, De Carli P et al (1997) Stridulatory organ and ultrasonic emission in certain species of Ponerine ants (Genus *Ectatomma* and *Pachycondyla*, Hymenoptera, Formicidae). Bioacoustics 8:209–221. https://doi.org/10.1080/09524622.1997.9753363
Paviotti M, Vogiatzis K (2012) On the outdoor annoyance from scooter and motorbike noise in the urban environment. Sci Total Environ 430:223–230. https://doi.org/10.1016/j.scitotenv.2012.05.010
Payne RS, McVay S (1971) Songs of humpback whales. Science 173:585–597. https://doi.org/10.1126/science.173.3997.585
Payne KB, Langbauer WR Jr, Thomas EM (1986) Infrasonic calls of the Asian elephant (*Elephas maximus*). Behav Ecol Sociobiol 18:297–301. https://doi.org/10.1007/BF00300007
Payne CJ, Jessop TS, Guay PJ et al (2012) Population, behavioural and physiological responses of an urban population of black swans to an intense annual noise event. PLoS One 7:e45014. https://doi.org/10.1371/journal.pone.0045014
Phillips HRP, Newbold T, Purvis A (2017) Land-use effects on local biodiversity in tropical forests vary between continents. Biodivers Conserv 26:2251–2270. https://doi.org/10.1007/s10531-017-1356-2
Picciulin M, Sebastianutto L, Fortuna CM et al (2016) Are the 1/3-octave band 63- and 125 Hz noise levels predictive of vessel activity? The case in the Cres-Lošinj Archipelago (northern Adreatic Sea, Croatia). In: Popper AN, Hawkins A (eds) The effects of noise on aquatic life II. Springer, New York, pp 821–828
Pieretti N, Farina A, Morri D (2011) A new methodology to infer the singing activity of an avian community: the Acoustic Complexity Index (ACI). Ecol Indic 11:868–873. https://doi.org/10.1016/j.ecolind.2010.11.005
Pijanowski BC, Farina A, Gage SH et al (2011a) What is soundscape ecology? An introduction and overview of an emerging new science. Landsc Ecol 26:1213–1232. https://doi.org/10.1007/s10980-011-9600-8
Pijanowski BC, Villanueva-Rivera LJ, Dumyahn SL et al (2011b) Soundscape ecology: the science of sound in the landscape. Bioscience 61:203–216. https://doi.org/10.1525/bio.2011.61.3.6
Pola YV, Snowdon CT (1975) The vocalizations of pygmy marmosets (*Cebuella pygmaea*). Anim Behav 23:826–842. https://doi.org/10.1016/0003-3472(75)90108-6
Polidori C, Pavan G, Ruffato G et al (2013) Common features and species-specific differences in stridulatory organs and stridulation patterns of velvet ants (Hymenoptera: Mutillidae). Zool Anz 252:457–468. https://doi.org/10.1016/j.jcz.2013.01.003
Potočnik I, Poje A (2010) Noise pollution in forest environment due to forest operations. Croat J For Eng 31:137–148
Prat Y, Taub M, Yovel Y (2016) Everyday bat vocalizations contain information about emitter, addressee, context, and behavior. Sci Rep 6:39419. https://doi.org/10.1038/srep39419
Prince P, Hill A, Covarrubias EP et al (2019) Deploying acoustic detection algorithms on low-cost, open-source sensors for environmental monitoring. Sensors 19:553. https://doi.org/10.3390/s19030553
Priyadarshani N, Castro I, Marsland S (2018) The impact of environmental factors in birdsong acquisition using automated recorders. Ecol Evol 8:5016–5033. https://doi.org/10.1002/ece3.3889
Putland RL, Mensinger AF (2020) Exploring the soundscape of small freshwater lakes. Ecol Inform 55:101018. https://doi.org/10.1016/j.ecoinf.2019.101018
Quadros S, Goulart VDL, Passos L et al (2014) Zoo visitor effect on mammal behaviour: does noise matter? Appl Anim Behav Sci 156:78–84. https://doi.org/10.1016/j.applanim.2014.04.002
Radford C, Jeffs A, Tindle C, Montgomery JC (2008) Resonating sea urchin skeletons create coastal choruses. Mar Ecol Prog Ser 362:37–43. https://doi.org/10.3354/meps07444
Rashed A, Khan MI, Dawson JW et al (2009) Do hoverflies (Diptera: Syrphidae) sound like the Hymenoptera they morphologically resemble? Behav Ecol 20:396–402. https://doi.org/10.1093/beheco/arn148
Reber SA, Janisch J, Torregrosa K et al (2017) Formants provide honest acoustic cues to body size in American alligators. Sci Rep 7:1816. https://doi.org/10.1038/s41598-017-01948-1
Reed SE, Boggs JL, Mann JP (2010) SPreAD-GIS: an ArcGIS toolbox for modeling the propagation of engine noise in a wildland setting. Version 2.0. The Wilderness Society, San Franscisco, CA
Reine KJ, Clarke DG, Dickerson C (2014) Characterization of underwater sounds produced by hydraulic and mechanical dredging operations. J Acoust Soc Am 135:3280–3294. https://doi.org/10.1121/1.4875712
Ricci SW, Eggleston DB, Bohnenstiehl DR, Lillis A (2016) Temporal soundscape patterns and processes in an estuarine reserve. Mar Ecol Prog Ser 550:25–38. https://doi.org/10.3354/meps11724
Rice AN, Soldevilla MS, Quinlan JA (2017) Nocturnal patterns in fish chorusing off the coasts of Georgia and eastern Florida. Bull Mar Sci 93:455–474. https://doi.org/10.5343/bms.2016.1043
Righini R, Pavan G (2020) First assessment of the soundscape of the Integral Nature Reserve "Sasso Fratino" in the Central Apennine, Italy. Biodiversity 21:4–14. https://doi.org/10.1080/14888386.2019.1696229
Roberts C (2009) Construction noise and vibration impact on sensitive premises. In: Proceedings of Acoustics 2009. Australian Acoustical Society, Adelaide, 23–25 November 2009
Robillard T, Montealegre-Z F, Desutter-Grandcolas L et al (2013) Mechanisms of high-frequency song generation in brachypterous crickets and the role of ghost

frequencies. J Exp Biol 216:2001–2011. https://doi.org/10.1242/jeb.083964
Rocha RC Jr, Clapham PJ, Ivashchenko YV (2014) Emptying the oceans: a summary of industrial whaling catches in the 20th century. Mar Fish Rev 76:37–48. https://doi.org/10.7755/MFR.76.4.3
Rochat JL, Reiter D (2016) Highway traffic noise. Acoust Today 12:38–47
Römer H, Lewald J (1992) High-frequency sound transmission in natural habitats: implications for the evolution of insect acoustic communication. Behav Ecol Sociobiol 29:437–444. https://doi.org/10.1007/BF00170174
Ross D (1976) Mechanics of underwater noise. Pergamon Press, Oxford
Rossi T, Connell SD, Nagelkerken I (2016a) Silent oceans: ocean acidification impoverishes natural soundscapes by altering sound production of the world's noisiest marine invertebrate. Proc R Soc B 283:20153046. https://doi.org/10.1098/rspb.2015.3046
Rossi T, Nagelkerken I, Pistevos JCA, Connell SD (2016b) Lost at sea: ocean acidification undermines larval fish orientation via altered hearing and marine soundscape modification. Biol Lett 12:20150937. https://doi.org/10.1098/rsbl.2015.0937
Rountree RA, Gilmore RG, Goudey CA et al (2006) Listening to fish: applications of passive acoustics to fisheries science. Fisheries 31:433–446. https://doi.org/10.1577/1548-8446(2006)31[433:LTF]2.0.CO;2
Rountree RA, Bolgan M, Juanes F (2019) How can we understand freshwater soundscapes without fish sound descriptions? Fisheries 44:137–143. https://doi.org/10.1002/fsh.10190
Sales GD, Wilson KJ, Spencer KEV (1988) Environmental ultrasound in laboratories and animal houses: a possible cause for concern in the welfare and use of laboratory animals. Lab Anim 22:369–375. https://doi.org/10.1258/002367788780746188
Salmi R, Hammerschmidt K, Doran-Sheehy DM (2013) Western gorilla vocal repertoire and contextual use of vocalizations. Ethology 119:831–847. https://doi.org/10.1111/eth.12122
Sáncez-Pérez LA, Sánchez-Fernández LP, Suárez-Guerra S, Carbajal-Hernández JJ (2013) Aircraft class identification based on take-off noise signal segmentation in time. Expert Syst Appl 40:5148–5159. https://doi.org/10.1016/j.eswa.2013.03.017
Scarpelli MDA, Ribeiro MC, Teixeira FZ et al (2020) Gaps in terrestrial soundscape research: it's time to focus on tropical wildlife. Sci Total Environ 707:135403. https://doi.org/10.1016/j.scitotenv.2019.135403
Schafer RM (1969) The new soundscape: a handbook for the modern music teacher. Associated Music Publishers, New York
Schafer RM (1977) The tuning of the world. Knopf, New York
Scheifele PM, Clark JG, Sonstrom K et al (2012a) Ballroom music spillover into a beluga whale aquarium exhibit. Adv Acoust Vib 2012:402130. https://doi.org/10.1155/2012/402130
Scheifele PM, Johnson MT, Kretschmer L et al (2012b) Ambient habitat noise and vibration at the Georgia aquarium. J Acoust Soc Am 132:EL88–EL94. https://doi.org/10.1121/1.4734387
Schmitz B (2002) Sound production in crustacea with special reference to the Alpheidae. In: Wiese K (ed) The crustacean nervous system. Springer, Berlin, pp 536–547
Schoeman RP, Erbe C, Plön S (2022) Underwater chatter for the win: a first assessment of underwater soundscapes in two bays along the Eastern Cape coast of South Africa. J Mar Sci Eng 10:746. https://doi.org/10.3390/jmse10060746
Schusterman RJ, Gentry R, Schmook J (1966) Underwater vocalizations by sea lions: social and mirror stimuli. Science 154:540–542. https://doi.org/10.1126/science.154.3748.540
Sèbe F, Aubin T, Boué A, Poindron P (2008) Mother-young vocal communication and acoustic recognition promote preferential nursing in sheep. J Exp Biol 211:3554–3562. https://doi.org/10.1242/jeb.016055
Sertlek HÖ, Slabbekoorn H, ten Cate C, Ainslie MA (2019) Source specific sound mapping: spatial, temporal and spectral distribution of sound in the Dutch North Sea. Environ Pollut 247:1143–1157. https://doi.org/10.1016/j.envpol.2019.01.119
Shannon G, McKenna MF, Angeloni LM et al (2016) A synthesis of two decades of research documenting the effects of noise on wildlife. Biol Rev 91:982–1005. https://doi.org/10.1111/brv.12207
Sherwen SL, Hemsworth PH (2019) The visitor effect on zoo animals: implications and opportunities for zoo animal welfare. Animals 9:366. https://doi.org/10.3390/ani9060366
Slabbekoorn H (2004) Habitat-dependent ambient noise: consistent spectral profiles in two African forest types. J Acoust Soc Am 116:3727–3733. https://doi.org/10.1121/1.1811121
Slabbekoorn H, Bouton N (2008) Soundscape orientation: a new field in need of sound investigation. Anim Behav 76:5–8. https://doi.org/10.1016/j.anbehav.2008.06.010
Slabbekoorn H, Smith TB (2002) Habitat-dependent song divergence in the little greenbul: an analysis of environmental selection pressures on acoustic signals. Evolution 56:1849–1858. https://doi.org/10.1111/j.0014-3820.2002.tb00199.x
Slabbekoorn H, Dooling RJ, Popper AN, Fay RR (2018) Effects of anthropogenic noise on animals. Springer, New York
Slabbekoorn H, Dalen J, de Haan D et al (2019) Population-level consequences of seismic surveys on fishes: an interdisciplinary challenge. Fish Fish 20:653–685. https://doi.org/10.1111/faf.12367
Soloway AG, Dahl PH (2014) Peak sound pressure and sound exposure level from underwater explosions in shallow water. J Acoust Soc Am 136:218–223. https://doi.org/10.1121/1.4892668

Soltis J (2010) Vocal communication in African elephants (*Loxodonta africana*). Zoo Biol 29:192–209. https://doi.org/10.1002/zoo.20251

Southworth M (1969) The sonic environment of cities. Environ Behav 1:49–70. https://doi.org/10.1177/001391656900100104

Staaterman ER, Claverie T, Patek SN (2010) Disentangling defense: the function of spiny lobster sounds. Behaviour 147:235–258. https://doi.org/10.1163/000579509X12523919243428

Stack DW, Peter N, Manning RE, Fristrup KM (2011) Reducing visitor noise levels at Muir Woods National Momument using experimental management. J Acoust Soc Am 129:1375–1380. https://doi.org/10.1121/1.3531803

Stirling I, Calvert W, Spencer C (1987) Evidence of stereotyped underwater vocalizations of male Atlantic walruses (*Odobenus rosmarus rosmarus*). Can J Zool 65:2311–2321. https://doi.org/10.1139/z87-348

Sueur J (2002) Cicada acoustic communication: potential sound partitioning in a multispecies community from Mexico (Hemiptera: Cicadomorpha: Cicadidae). Biol J Linn Soc 75:379–394. https://doi.org/10.1046/j.1095-8312.2002.00030.x

Sueur J (2018) Sound analysis and synthesis with R. Springer, Cham

Sueur J, Aubin T, Simonis C (2008a) "Seewave": a free modular tool for sound analysis and synthesis. Bioacoustics 18:213–226. https://doi.org/10.1080/09524622.2008.9753600

Sueur J, Pavoine S, Hamerlynck O, Duvail S (2008b) Rapid acoustic survey for biodiversity appraisal. PLoS One 3:e4065. https://doi.org/10.1371/journal.pone.0004065

Sueur J, Krause B, Farina A (2019) Climate change is breaking earth's beat. Trends Ecol Evol 34:971–973. https://doi.org/10.1016/j.tree.2019.07.014

Talandier J, Hyvernaud O, Reymond D, Okal EA (2006) Hydroacoustic signals generated by parked and drifting icebergs in the southern Indian and Pacific oceans. Geophys J Int 165:817–834. https://doi.org/10.1111/j.1365-246X.2006.02911.x

Tasker ML, Amundin M, Andre M, et al (2010) Marine Strategy Framework Directive - Task Group 11: underwater noise and other forms of energy. JRC Scientific and Technical Report EUR 24341 EN - 2010. Office for Official Publications of the European Communities, Luxembourg

Tennessen JB, Parks SE (2016) Acoustic propagation modeling indicates vocal compensation in noise improves communication range for North Atlantic right whales. Endanger Species Res 30:225–237. https://doi.org/10.3354/esr00738

Tepp G, Chadwick WW Jr, Haney MM et al (2019) Hydroacoustic, seismic, and bathymetric observations of the 2014 submarine eruption at Ahyi seamount, Mariana arc. Geochem Geophys Geosyst 20:3608–3627. https://doi.org/10.1029/2019GC008311

Thiebault A, Charrier I, Aubin T et al (2019) First evidence of underwater vocalisations in hunting penguins. PeerJ 7:e8340. https://doi.org/10.7717/peerj.8240

Thomas JA, Kuechle VB (1982) Quantitative analysis of Weddell seal (*Leptonychotes weddelli*) underwater vocalizations at McMurdo Sound, Antarctica. J Acoust Soc Am 72:1730–1738. https://doi.org/10.1121/1.388667

Tonolla D, Acuña V, Lorang MS et al (2010) A field-based investigation to examine underwater soundscapes of five common river habitats. Hydrol Process 24:3146–3156. https://doi.org/10.1002/hyp.7730

Tonolla D, Lorang MS, Heutschi K et al (2011) Characterization of spatial heterogeneity in underwater soundscapes at the river segment scale. Limnol Oceanogr 56:2319–2333. https://doi.org/10.4319/lo.2011.56.6.2319

Torigoe K (1982) A study of the world soundscape project. York University, Toronto, Ontario

Tougaard J, Henriksen OD, Miller LA (2009) Underwater noise from three types of offshore wind turbines: estimation of impact zones for harbor porpoises and harbor seals. J Acoust Soc Am 125:3766–3773. https://doi.org/10.1121/1.3117444

Towsey M, Wimmer J, Williamson I, Roe P (2014) The use of acoustic indices to determine avian species richness in audio-recordings of the environment. Ecol Inform 21:110–119. https://doi.org/10.1016/j.ecoinf.2013.11.007

Tripovich JS, Hall-Aspland S, Charrier I, Arnould JPY (2012) The behavioural response of Australian fur seals to motor boat noise. PLoS One 7:e37228. https://doi.org/10.1371/journal.pone.0037228

Truax B (1984) Acoustic communication. Ablex Publishing Corporation, Norwood

Truax B (1996) Soundscape, acoustic communication and environmental sound composition. Contemp Music Rev 15:49–65

van der Lee GH, Desjonquères C, Sueur J, Kraak MHS, Verdonschot PFM (2020) Freshwater ecoacoustics: Listening to the ecological status of multi-stressed lowland waters. Ecol Indic 113:106252. https://doi.org/10.1016/j.ecolind.2020.106252

van Opzeeland I, van Parijs S, Bornemann H et al (2010) Acoustic ecology of Antarctic pinnipeds. Mar Ecol Prog Ser 414:267–291. https://doi.org/10.3354/meps08683

Van Parijs SM, Kovacs KM (2002) In-air and underwater vocalizations of eastern Canadian harbour seals, *Phoca vitulina*. Can J Zool 80:1173–1179. https://doi.org/10.1139/z02-088

Veirs S, Veirs V, Wood JD (2016) Ship noise extends to frequencies used for echolocation by endangered killer whales. PeerJ 4:e1657. https://doi.org/10.7717/peerj.1657

Vergne AL, Pritz MB, Mathevon N (2009) Acoustic communication in crocodilians: from behaviour to brain. Biol Rev 84:391–411. https://doi.org/10.1111/j.1469-185X.2009.00079.x

Vergne AL, Aubin T, Taylor P, Mathevon N (2011) Acoustic signals of baby black caimans. Zoology 114:313–320. https://doi.org/10.1016/j.zool.2011.07.003

Vidović A, Štimac I, Zečević-Tadić R (2017) Aircraft noise monitoring in function of flight safety and aircraft model determination. J Adv Transp 2017: 2850860. https://doi.org/10.1155/2017/2850860

Villanueva-Rivera LJ (2014) *Eleutherodactylus* frogs show frequency but no temporal partitioning: implications for the acoustic niche hypothesis. PeerJ 2:e496. https://doi.org/10.7717/peerj.496

Villanueva-Rivera LJ, Pijanowski BC (2018) "Soundecology": soundscape ecology. R package version 1.3.3

Villanueva-Rivera LJ, Pijanowski BC, Doucette J, Pekin B (2011) A primer of acoustic analysis for landscape ecologists. Landsc Ecol 26:1233–1246. https://doi.org/10.1007/s10980-011-9636-9

von Benda-Beckmann AM, Aarts G, Sertlek HO et al (2015) Assessing the impact of underwater clearance of unexploded ordnance on harbour porpoises (*Phocoena phocoena*) in the southern North Sea. Aquat Mamm 41:503–523. https://doi.org/10.1578/AM.41.4.2015.503

Walker SE, Cade WH (2003) The effects of temperature and age on calling song in a field cricket with a complex calling song, *Teleogryllus oceanicus* (Orthoptera: Gryllidae). Can J Zool 81:1414–1420. https://doi.org/10.1139/z03-106

Wark JD (2015) The influence of the sound environment on the welfare of zoo-housed callitrichine monkeys. Case Western Reserve University, Cleveland

Weilgart L, Whitehead H (1993) Coda communication by sperm whales (*Physeter macrocephalus*) off the Galápagos Islands. Can J Zool 71:744–752. https://doi.org/10.1139/z93-098

Wellard R, Pitman RL, Durban J, Erbe C (2020) Cold call: the acoustic repertoire of Ross Sea killer whales (*Orcinus orca*, Type C) in McMurdo Sound, Antarctica. R Soc Open Sci 7:191228. https://doi.org/10.1098/rsos.191228

Wemmer C, von Ebers M, Scow K (1976) An analysis of the chuffing vocalization in the polar bear (*Ursus maritimus*). J Zool 180:425–439. https://doi.org/10.1111/j.1469-7998.1976.tb04686.x

Wenz GM (1962) Acoustic ambient noise in the ocean: spectra and sources. J Acoust Soc Am 34:1936–1956. https://doi.org/10.1121/1.1909155

White K, Arntzen M, Walker F et al (2017) Noise annoyance caused by continuous descent approaches compared to regular descent procedures. Appl Acoust 125: 194–198. https://doi.org/10.1016/j.apacoust.2017.04.008

Williams R, Erbe C, Ashe E, Clark CW (2015) Quiet(er) marine protected areas. Mar Pollut Bull 100:154–161. https://doi.org/10.1016/j.marpolbul.2015.09.012

Wiseman S, Wilson PS, Sepulveda F (2014) Measuring a soundscape of the captive southern white rhinoceros (*Ceratotherium simum simum*). In: Davy J, Don C, McMinn T et al (eds) Proceedings of 43rd International Congress on Noise Control Engineering. The Australian Acoustical Society, Melbourne, 16–19 November 2014

Wolfenden AD, Slabbekoorn H, Kluk K, de Kort SR (2019) Aircraft sound exposure leads to song frequency decline and elevated aggression in wild chiffchaffs. J Anim Ecol 88:1720–1731. https://doi.org/10.1111/1365-2656.13059

Wrege PH, Rowland ED, Keen S, Shiu Y (2017) Acoustic monitoring for conservation in tropical forests: examples from forest elephants. Methods Ecol Evol 8:1292–1301. https://doi.org/10.1111/2041-210X.12730

Wyatt R (2008) Review of existing data on underwater sounds produced by the oil and gas industry. Joint Industry Programme on Sound and Aquatic Life, London

Wysocki LE, Davidson JW III, Smith ME et al (2007) Effects of aquaculture production noise on hearing, growth, and disease resistance of rainbow trout *Oncorhynchus mykiss*. Aquaculture 272:687–697. https://doi.org/10.1016/j.aquaculture.2007.07.225

Xu W, Dong L, Caruso F, Gong Z, Li S (2020) Long-term and large-scale spatiotemporal patterns of soundscape in a tropical habitat of the Indo-Pacific humpback dolphin (*Sousa chinensis*). PLoS One 15:e0236938. https://doi.org/10.1371/journal.pone.0236938

Yin S, McCowan B (2004) Barking in domestic dogs: context specificity and individual identification. Anim Behav 68:343–355. https://doi.org/10.1016/j.anbehav.2003.07.016

Yip DA, Bayne EM, Sólymos P et al (2017) Sound attenuation in forest and roadside environments: implications for avian point-count surveys. Condor Ornithol Appl 119:73–84. https://doi.org/10.1650/CONDOR-16-93.1

York D (1994) Recreational-boating disturbances of natural communities and wildlife: an annotated bibliography. Biological Report 22. U.S. Department of the Interior, Washington

Young BA (1991) Morphological basis of "growling" in the king cobra, *Ophiophagus hannah*. J Exp Zool 260: 275–287. https://doi.org/10.1002/jez.1402600302

Young BA (2003) Snake bioacoustics: toward a richer understanding of the behavioral ecology of snakes. Q Rev Biol 78:303–325. https://doi.org/10.1086/377052

Young BA, Brown IP (1993) On the acoustic profile of the rattlesnake rattle. Amphibia Reptilia 14:373–380. https://doi.org/10.1163/156853893X00066

Zhang F, Zhao J, Feng AS (2017) Vocalizations of female frogs contain nonlinear characteristics and individual signatures. PLoS One 12:e0174815. https://doi.org/10.1371/journal.pone.0174815

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



# Detection and Classification Methods for Animal Sounds

8

Julie N. Oswald, Christine Erbe, William L. Gannon, Shyam Madhusudhana, and Jeanette A. Thomas

## 8.1 Introduction

Researchers have a natural tendency to classify biological systems into categories. For example, organisms can be classified based on biome, ecosystem, taxon, phylogeny, niche, demographic class, behavior type, etc., and this allows complex systems to be organized. Categorization also can make recognition of patterns easier and assist in understanding the ways in which biological systems work. Classification provides a convenient method for comparing features, making systematic measurements, testing hypotheses, and performing statistical analyses.

Bioacousticians have categorized sounds produced by animals for decades, and new methods for classification continue to be developed (Horn and Falls 1996; Beeman 1998). Animals produce many different types of sounds that span orders of magnitude along the dimensions of time, frequency, and amplitude. For example, the repertoire of marine mammal acoustic signals includes broadband echolocation clicks as short as 10 μs in duration and with energy up to 200 kHz, as well as narrowband tonal sounds as low as 10–20 Hz, lasting more than10 s. Song birds and some species of baleen whales arrange individual sounds into patterns called song and repeat these patterns for hours or days. Some mammal species produce distinctive, stereotyped sounds (e.g., chipmunks, dogs, and blue whales), while others produce signals with high variability (e.g., mimicking birds, primates, and dolphins).

Because animals produce so many different types of sounds, developing algorithms to detect, recognize, and classify a wide range of acoustic signals can be challenging. In the past, detection and classification tasks were performed by an experienced bioacoustician who listened to the sounds and visually reviewed spectrographic displays (e.g., for birds by Baptista and Gaunt 1997; chipmunks by Gannon and Lawlor 1989; baleen whales by Stafford et al. 1999; and delphinids by Oswald et al. 2003). Before the advent of digital signal-analysis, data were

Jeanette A. Thomas (deceased) contributed to this chapter while at the Department of Biological Sciences, Western Illinois University-Quad Cities, Moline, IL, USA

J. N. Oswald (✉)
Scottish Oceans Institute, University of St Andrews, St Andrews, Fife, UK
e-mail: jno@st-andrews.ac.uk

C. Erbe
Centre for Marine Science & Technology, Curtin University, Perth, WA, Australia
e-mail: c.erbe@curtin.edu.au

W. L. Gannon
Department of Biology and Graduate Studies, Museum of Southwestern Biology, University of New Mexico, Albuquerque, NM, USA
e-mail: wgannon@unm.edu

S. Madhusudhana
K. Lisa Yang Center for Conservation Bioacoustics, Cornell Lab of Ornithology, Cornell University, Ithaca, NY, USA
e-mail: shyamm@cornell.edu

© The Author(s) 2022
C. Erbe, J. A. Thomas (eds.), *Exploring Animal Behavior Through Sound: Volume 1*,
https://doi.org/10.1007/978-3-030-97540-1_8

analyzed while enduring the acrid smell of etched Kay Sona-Graph paper and piles of 8-s printouts removed from a spinning recording drum littering laboratory tables and floors. Output from a long-duration sound had to be spliced together (see Chap. 1). Many bioacoustic studies generated an enormous amount of data, which made this manual review process at best inefficient, and at worst impossible to accomplish.

For decades, scientists have worked to automate the process of detecting and classifying sounds into categories or types. Automated classification involves three main steps: (1) detection of potential sounds of interest, (2) extraction of relevant acoustic characteristics (or, features) from these sounds, and (3) classification of these sounds as produced by a particular species, sex, age, or individual. Methods for the automated detection of sounds have progressed quickly with technological advances in digital recording (see Chap. 2). Likewise, the extraction of sound variables useful in analysis has expanded with an increasing amount of information provided by new technology. For instance, where features such as maximum frequency or time between sounds originally were measured manually off sonagraph paper, devices today allow for measuring these, and many more variables, automatically or semi-automatically using computer software. Now, derived variables, such as time difference between individual signal elements, frequency modulation, running averages of sound frequency, and harmonic structure can be easily obtained for classifying the sounds in a repertoire.

Some of the earliest methods used for automated detection and classification included energy threshold detectors (e.g., Clark 1980) and matched filters (e.g., Freitag and Tyack 1993; Stafford et al. 1998; Dang et al. 2008; Mankin et al. 2008). These methods were used to detect and classify simple, stereotypical sounds produced by species such as the Asian longhorn beetle (*Anoplophora glabripennis*), cane toads (*Rhinella marina*), blue whales (*Balaenoptera* spp.), and fin whales (*Balaenoptera physalus*). Once sounds are detected, they can be organized into groups, or classified, based on selected acoustic characteristics. For example, development of methods for detection and automated signal processing of bat sounds led to a variety of automated, off-the-shelf, ready-to-deploy bat detectors that detect and classify sounds by species (Fenton and Jacobson 1973; Gannon et al. 2004). These detectors can be very useful in addressing biological or management issues in ecology, evolution, and impact mitigation. While the accuracy and robustness of automated approaches are always a matter of concern (Herr et al. 1997; Parsons et al. 2000), modern techniques promise much improved recognition performances that could rival manual analyses (e.g., Brown and Smaragdis 2009).

Multivariate statistical methods can be powerful for classification of sounds produced by species with variable vocal repertoires because they can identify complex relationships among many acoustic features (see Chap. 9). With the advent of powerful personal computers in the 1980s and 1990s, the use of multivariate techniques became popular for classifying bird sounds (e.g., Sparling and Williams 1978; Martindale 1980a, b). Since then, enormous effort has been expended to develop these and other automatic methods for the detection of sounds produced by many taxa and their classification into discrete categories, such as species, population, sex, or individual.

These days, there are applications (apps) for smartphones that use advanced algorithms to automatically detect and recognize sounds. For example, the *BirdNET* app detects and classifies bird song—similar to the *Shazam* app for music—and provides a listing of the top-ranked matching species. It includes almost 1000 of the most common species of North America and Europe. A similar app, *Song Sleuth*, recognizes songs of nearly 200 bird species likely to be heard in North America and also provides references for species identification, such as the David Sibley Bird Reference (Sibley 2000), allowing the user to "dig into" the bird's biology and conservation needs.

In this chapter, we present an overview of methods for detection and classification of sounds along with examples from different taxa. No single method is appropriate for every research

project and so the strengths and weaknesses of each method are summarized to help guide decisions on which methods are better suited for particular research scenarios. Because algorithms for statistical analyses, automated detection, and computer classification of animal sounds are advancing rapidly, this is not a comprehensive overview of methods, but rather a starting point to stimulate further investigations.

## 8.2 Qualitative Naming and Classification of Animal Sounds

Prior to computer-assisted detection and classification of animal sounds, bioacousticians used various qualitative methods to categorize sounds.

### 8.2.1 Onomatopoeic Names

Frequently, researchers describe and name animal sounds based on their perception of the sound and thus based on their own language. This approach has been common in the study of terrestrial animals (in particular, birds) and marine mammals (in particular, pinnipeds and mysticetes). Researchers also have given onomatopoeic names to sounds. These are names that phonetically resemble the sound they describe. For example, the sounds of squirrels and chipmunks have been described as barks, chatters, chirps, and growls. The primate literature is also rich in these sorts of sound descriptions (e.g., the hack sequences and boom-hack sequences described for Campbell's monkeys, *Cercopithecus campbelli*; Ouattara et al. 2009). Bioacousticians studying humpback whales (*Megaptera novaeangliae*) have described a repertoire of sounds including barks, bellows, chirps, cries, croaks, groans, growls, grumbles, horns, moans, purrs, screams, shrieks, sighs, sirens, snorts, squeaks, thwops, trumpets, violins, wops, and yaps (Dunlop et al. 2007, 2013). While it is potentially convenient for researchers within a group to discuss sounds this way, it is more difficult for others, and perhaps impossible for foreign-language speakers to recognize the sound type. An example of this difficulty in describing a sound is the ubiquitous rooster crow, which can be described by a US citizen as "cock-a-doodle-doo" and by a German citizen as "kikeriki". Roosters make the same sound, no matter in which country they live, yet their single sound has been named so differently, as has the bark of dogs (Fig. 8.1). Of course, onomatopoeic naming of sounds also fails when the sounds are outside of the human hearing range.

If the above was not confusing enough, bird calls have been described using onomatopoeic phrases. For example, the song of a white-throated sparrow (*Zonotrichia albicollis*) has been described in Canada as sounding like "O sweet Canada Canada Canada" and in New England, USA, as "Old Sam Peabody Peabody Peabody." Another example is the barred owl (*Strix varia*), which hoots "Who cooks for you? Who cooks for you all?".

### 8.2.2 Naming Sounds Based on Animal Behavior

Researchers sometimes name sounds based on observed and interpreted animal behavior. For example, the various echolocation signals described for insectivorous bats have been named "search clicks" (i.e., slow and regular clicks) while pursuing insect prey and "terminal feeding buzz" (i.e., accelerated click trains) during prey capture (Griffin et al. 1960). The bird and mammal literature is replete with sounds named for a behavior, such as the begging call of nestling chicks (Briskie et al. 1999; Leonard and Horn 2001), the contact call for isolated young (Kondo and Watanabe 2009), and the alarm call warning of a nearby predator (Zuberbuhler et al. 1999; Gill and Bierema 2013). In some cases, the function of sounds has been studied in detail, which justifies using their function in the name. Examples are feeding buzzes in echolocation or alarm calls in primates. However, naming sounds according to behavior can be misleading because a sound can be associated with several contexts. Names based on the associated behavior should really only be used after detailed studies of context-specificity of the calls in question.

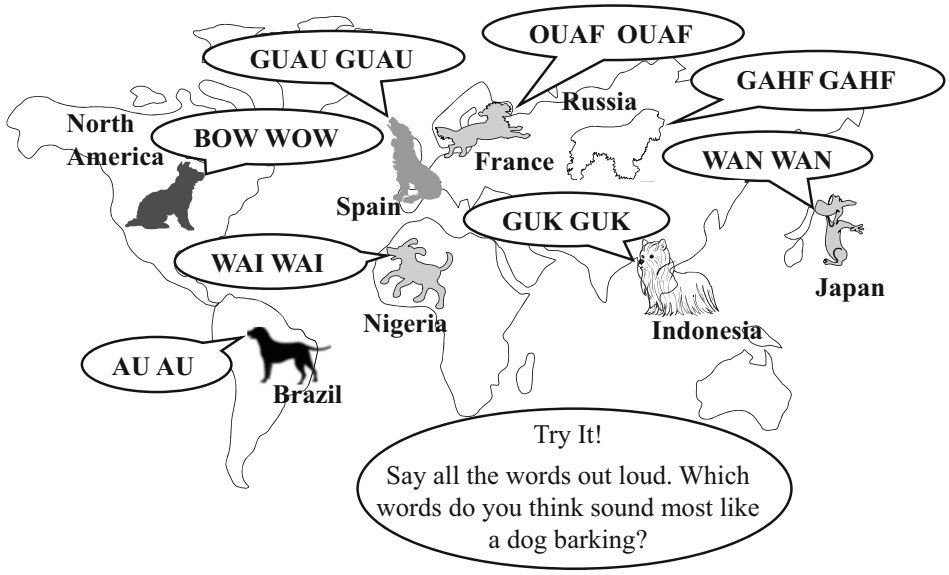



**Fig. 8.1** Dogs speak out. Labels used for dog barks in different countries

### 8.2.3 Naming Sounds Based on Mechanism of Sound Production

Some bioacousticians identify and classify sounds based on the mechanism of sound production. For example, one syllable in insect song corresponds to a single to- and fro-movement of a stridulatory anatomy or one cycle of a forewing opening and closing in the field cricket (*Gryllus* spp.). McLister et al. (1995) defined a note in chorusing frogs as the sound unit produced during a single expiration. Classifying sound types by their mode of production perhaps is less ambiguous and unequivocal, but there are limited data on the mechanisms of sound production in many animals.

### 8.2.4 Naming Sounds Based on Spectro-Temporal Features

An alternative, but not necessarily better, way of naming sounds is based on their spectro-temporal features. For instance, in distinguishing two morphologically similar species of bats, *Myotis californicus* is referred to as a "50-kHz bat" and *M. ciliolabrum* as a "40-kHz bat," which describes the terminal frequency of the downsweep of their ultrasonic echolocation signals (Gannon et al. 2001). Under water, the most common sound recorded from southern right whales (*Eubalaena australis*) is a 1–2 s frequency-modulated (FM) upsweep from about 50–200 Hz, commonly recorded with overtones, and referred to in the literature as the upcall (Fig. 8.2; Clark 1982). Antarctic blue whales (*Balaenoptera musculus intermedia*) produce a Z-call, which consists of a 10-s constant frequency (also called constant-wave, CW) sound at 28 Hz, followed by a rapid FM downsweep to 18 Hz, where the sound continues for another 15-s CW component (Rankin et al. 2005).

While the measurement of features from spectrograms and waveforms can be expected to be more objective than onomatopoeic or functional naming, the appearance of a spectrogram, and thus the measurements made, depend on characteristics of the recording system, the time and frequency settings of the analysis algorithm, and analysis algorithm used. This can make sounds look rather different at various scales and therefore lead to inconsistent classification.

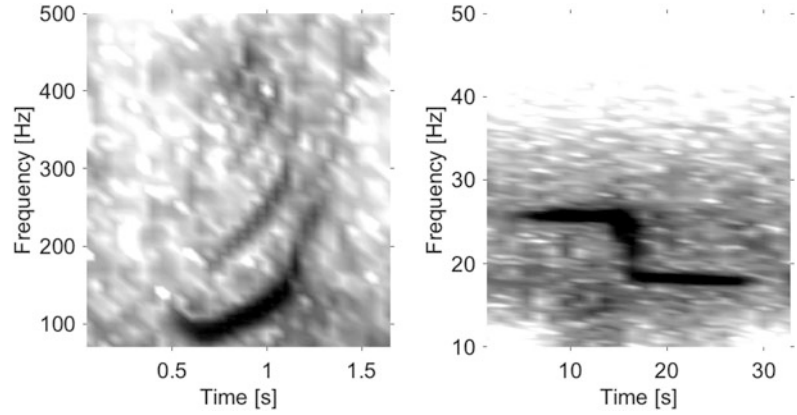



**Fig. 8.2** Spectrograms of southern right whale "upcall" (left; sampling frequency $f_s = 12$ kHz, Fourier window length $NFFT = 1200$, 50% overlap, Hann window) and Antarctic blue whale "Z-call" (right; $f_s = 6$ kHz, $NFFT = 16384$, 50% overlap, Hann window) recorded off southern Australia (Erbe et al. 2017)

An example of the confusion that can arise from different representations of sound is the boing sound made by minke whales (*Balaenoptera acutorostrata*), which was given an onomatopoeic name. In spectrograms, the boing might look like an FM sound (Fig. 8.3a), however, it is actually a series of rapid pulses (Rankin and Barlow 2005), similar to burst-pulse sounds produced by odontocetes (e.g., Wellard et al. 2015). As another example, the bioduck sound made by Antarctic minke whales (*Balaenoptera bonaerensis*) got its name because it resembles a duck's quack to human listeners (Risch et al. 2014). A spectrogram of the bioduck sound appears as a series of pulses; however, each pulse actually is a 0.3-s FM downswept tone from 300 to 100 Hz (Fig. 8.3b). As if this was not enough in terms of interesting sounds and odd names, dwarf minke whales produce the so-called star-wars sound, which is composed of a series of pulses with varying pulse rates (Gedamke et al. 2001). The different pulse rates make this sound appear as a mixture of broadband pulses and FM sounds in spectrograms, depending on the spectrogram settings (Fig. 8.3c). The sound name presumes the reader is familiar with the soundtrack of an American movie from the 1970s.

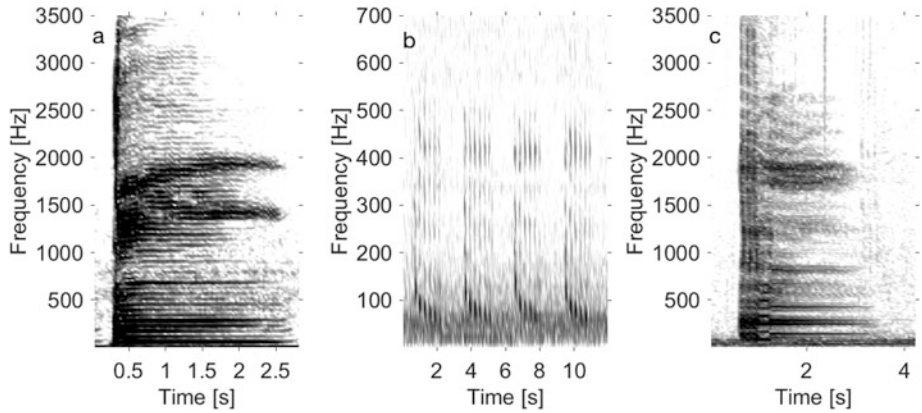



**Fig. 8.3** Spectrograms of the dwarf minke whale boing (**a** $f_s = 16$ kHz, $NFFT = 1024$, 50% overlap, Hann window), the Antarctic minke whale bioduck sound (**b** $f_s = 96$ kHz, $NFFT = 8192$, 50% overlap, Hann window), and the dwarf minke whale star-wars sound (**c** $f_s = 44$ kHz, $NFFT = 4096$, 50% overlap, Hann window). Recordings **a** and **b** from Erbe et al. (2017), **c** from Gedamke et al. (2001)

### 8.2.5 Naming Sounds Based on Human Communication Patterns

The term "song" is perhaps the best-known example of using human communication labels in the description of animal sounds. The word "song" may be used to simply indicate long-duration displays of a specific structure. Songs of insects and frogs are relatively simple sequences, consisting of the same sound repeated over long periods of time. The New River tree frog (*Trachycephalus hadroceps*), for example, produces nearly 38,000 calls in a single night (Starnberger et al. 2014). Many frogs use trilling notes in mate attraction, which has been described as song, but switch to a different vocal pattern in aggressive territorial displays (Wells 2007). In some frog songs, different notes serve different purposes, with one type of note warding off competing males, and another attracting females. In birds and mammals, songs are often more complex, consisting of several successive sounds in a recognizable pattern. They appear to be used primarily for territorial defense or mate attraction (Bradbury and Vehrencamp 2011). Our statements in this chapter show one way to describe calls and songs in animals; however, it is important to note that borrowing terminology from human communication when studying animals can lead to confusion. The terms we discuss here are not well defined and are used differently by different authors. Make sure to pay close attention to these definitions when reading literature about animal communication.

Some ornithologists have used human-language properties further to describe the structure of bird song. Song may be broken down into phrases (also called motifs). Each phrase is composed of syllables, which consist of notes (or elements, the smallest building blocks; Catchpole and Slater 2008). Notes, syllables, and phrases are identified and defined based on their repeated occurrence. An entire taxon of birds (songbirds, Order Passeriformes) has been designated by ornithologists because of their use of these elaborate sounds for territorial defense and/or mate attraction. Birds of this taxon usually use sets of sounds that are repeated in an organized structure. In many species, males produce such songs continuously for several hours each day, producing thousands of songs in each performance. In the bird song literature, songs are distinguished from calls by their more complex and sustained nature, species-typical patterns, or syntax that governs their combination of syllables and notes into a song. Songs are under the influence of reproductive hormones and associated with courtship (Bradbury and Vehrencamp 2011). Bird song can vary geographically and over time (e.g., Fig. 8.4; Camacho-Alpizar et al. 2018). In contrast, calls are typically acoustically simple and serve non-reproductive, maintenance functions, such as coordination of parental duties, foraging, responding to threats of predation, or keeping members of a group in contact (Marler 2004).

Several terrestrial mammals have been reported to sing. For instance, adult male rock hyraxes (*Procavia capensis*) engage throughout most of the year in rich and complex vocalization behavior that is termed singing (Koren et al. 2008). These songs are complex signals and are composed of multiple elements (chucks, snorts, squeaks, tweets, and wails) that encode the identity, age, body mass, size, social rank, and hormonal status of the singer (Koren and Geffen 2009, 2011). Holy and Guo (2005) described ultrasonic sounds from male laboratory mice (*Mus musculus*) as song. Von Muggenthaler et al. (2003) reported that Sumatran rhinoceros (*Dicerorhinus sumatrensis*) produce a song composed of three sound types: eeps (simple short signals, 70 Hz–4 kHz), humpback whale like sounds (100 Hz–3.2 kHz, varying in length, only produced by females), and whistle blows (loud, 17 Hz–8 kHz vocalizations followed by a burst of air with strong infrasonic content). Clarke et al. (2006) described the syntax and meaning of wild white-handed gibbon (*Hylobates lar*) songs.

Among marine mammals, blue, bowhead (*Balaena mysticetus*), fin, humpback, minke, and right whales, Weddell seals (*Leptonychotes weddellii*), harbor seals (*Phoca vitulina*), and

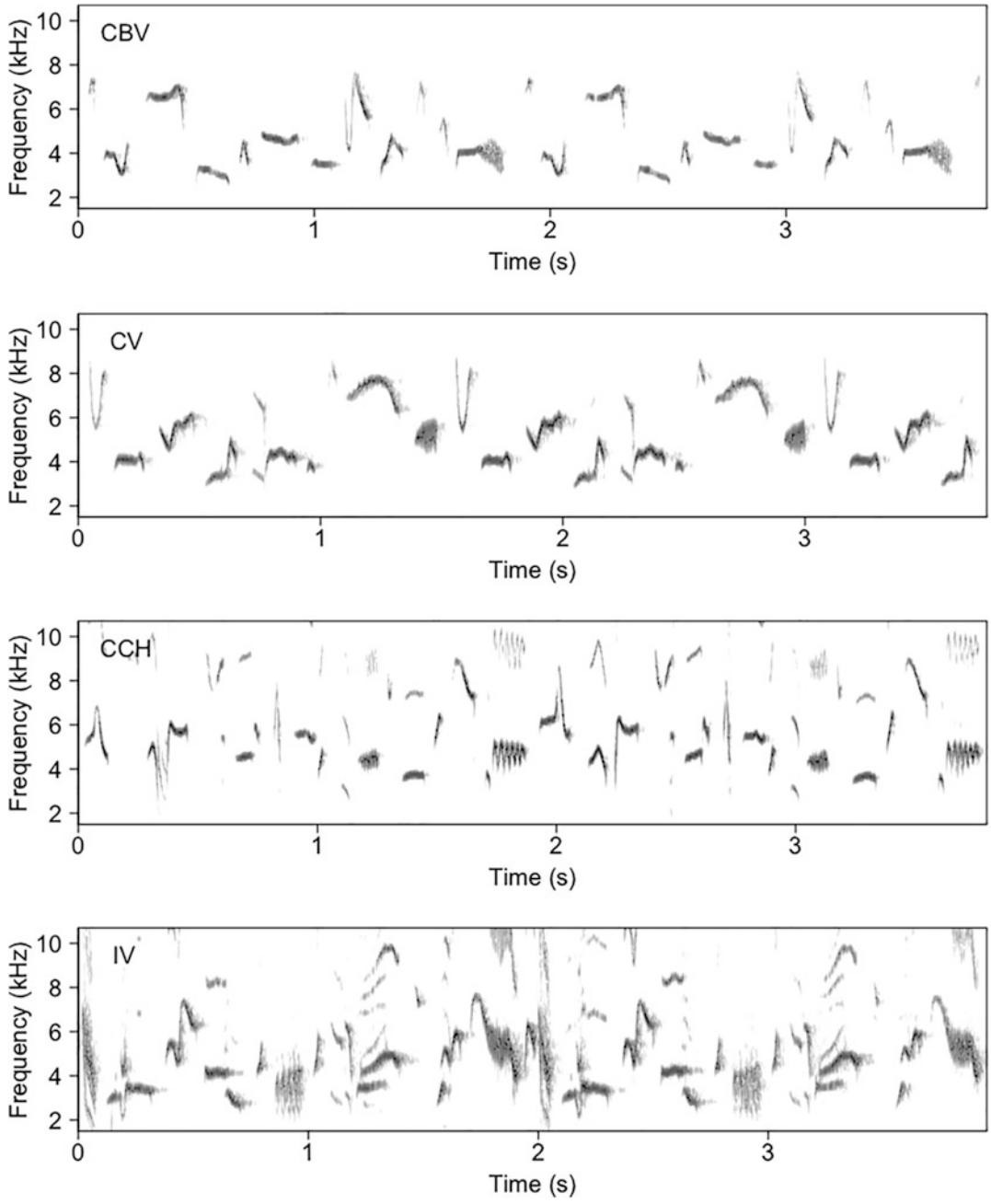



**Fig. 8.4** Geographic variation in birdsong. These spectrograms show a portion of song from Timberline wrens (*Thryorchilus browni*) recorded at four locations in Costa Rica (CBV = Cerro Buena Vista, CV = Cerro Vueltas, CCH = Cerro Chirripó, IV = Irazú Volcano) (Camacho-Alpizar et al. 2018). © Camacho-Alpizar et al.; https://doi.org/10.1371/journal.pone.0209508. Licensed under CC BY 4.0; https://creativecommons.org/licenses/by/4.0/

walrus (*Odobenus rosmarus*) have all been reported to sing (Payne and Payne 1985; Sjare et al. 2003; McDonald et al. 2006; Stafford et al. 2008; Oleson et al. 2014; Crance et al. 2019). The songs of blue, bowhead, fin, minke, and right whales are simple compared to those of the humpback whale and little is known about the behavioral context of song in any marine mammal species besides the humpback whale. Humpback whales are well-known for their long, elaborate songs. These songs are composed of themes consisting of repetitions of phrases made up of patterns of units similar to syllables in bird song (Fig. 8.5; Payne and Payne 1985; Helweg et al. 1998). Winn and Winn (1978) suggested that only male baleen whales sing, as a means of reproductive display. Sjare et al. (2003) reported that Atlantic walrus produce two main songs: the coda song and the diving vocalization song that differ by their pattern of knocks, taps, and bell sounds.

Song production does not exclude the emission of non-song sounds and most singing species likely emit both. The non-song sounds of humpback and pygmy blue whales (*Balaenoptera musculus brevicauda*), for example, have been cataloged (e.g., Recalde-Salas et al. 2014, 2020). Some song units may resemble non-song sounds.

Whether sounds are part of song or not, their detection and classification can be challenging when repertoires are large and possibly variable across time and space. Humpback whale songs, for example, vary by region and year (Cerchio et al. 2001; Payne and Payne 1985). Characterizing and describing the structure of song can be a difficult task even for the experienced bioacoustician. With the assistance of computer analysis tools, sound detection and classification may be more efficient.

## 8.3 Detection of Animal Sounds

The problem to be solved may seem simple. For example, a bioacoustician deployed an autonomous recorder in the field for a month, and after recovery of the gear, downloaded all data in the laboratory and now wants to pick all frog calls recorded in order to study the mating behavior of this species. Listening to the first few minutes of recording, the bioacoustician can easily hear the target species, but there are calls every few seconds—too many to pick by hand. So, the scientist looks for software tools to help detect all frog signals, and potentially sort them based on their acoustic features. The first step, signal detection, is discussed in Sect. 8.3; the second step, signal classification, is discussed in Sect. 8.4.

Automated signal detectors work by common principles. The raw input data are the ideally calibrated time series of pressure recorded with a microphone in air or hydrophone in water. There might be one or more pre-processing steps to filter or Fourier transform the data in successive time windows (see Chap. 4). The pre-processed time series is then fed into the detector, which computes a specific quantity from the acoustic data. This may be instantaneous energy, energy within a specified time window, entropy, or a correlation coefficient, as a few examples. Then, a detection threshold is applied. If the quantity exceeds the threshold, the signal is deemed present, otherwise not.

The threshold is commonly computed the following way:

$$E_{\text{th}} = \bar{E} + \gamma \sigma_E$$

where $E$ symbolizes the chosen quantity (e.g., energy), $\bar{E}$ is its mean value computed over a long time window (e.g., an entire file), $\sigma_E$ is the standard deviation, and $\gamma$ is a multiplier (integer or real). Setting a high threshold will result in only the strongest signals being detected and weaker ones being missed. Setting a low threshold will result in many false alarms, which are not signals. By varying $\gamma$, the ideal threshold may be found and the performance of the detector may be assessed (see Sect. 8.3.6).

### 8.3.1 Energy Threshold Detector

One of the most common methods for detecting animal sounds from recordings is to measure the

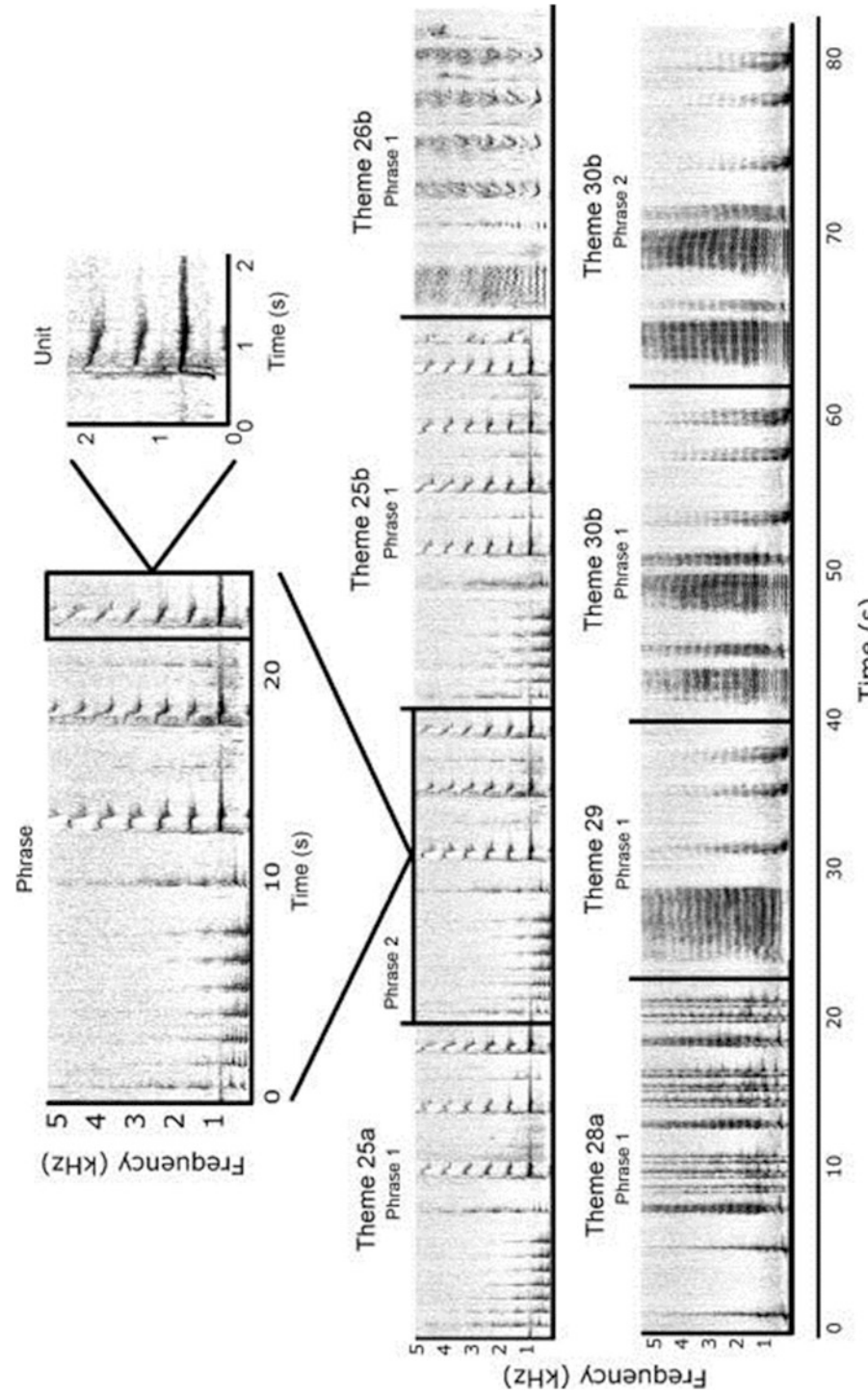



**Fig. 8.5** Spectrogram of the song structure of humpback whales, with sounds organized by theme, phrases, and units (Garland et al. 2017). © Acoustical Society of America, 2017. All rights reserved

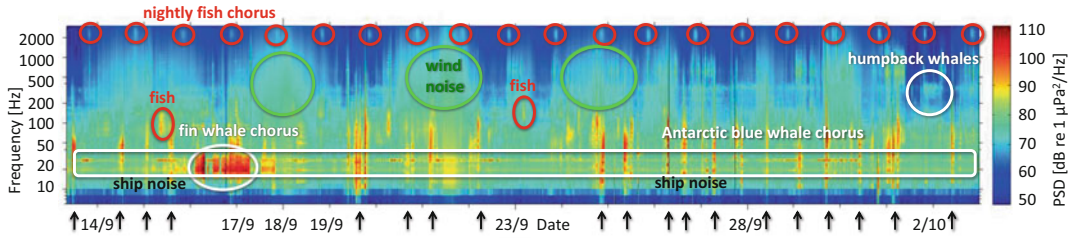



**Fig. 8.6** Spectrogram showing three weeks of choruses by fish, fin whales, and blue whales in the Perth Canyon, Australia (modified from Erbe et al. 2015). Fish raised ambient levels by 20 dB in the 1800–2500 Hz band every night. Fin whales raised ambient levels by 20 dB in the 15–35 Hz band over two days. Antarctic blue whales were the cause of ongoing tones at 18 and 28 Hz for weeks at a time. Colors represent power spectral density (PSD). Black arrows point to strong noise from passing ships. © Erbe et al.; https://doi.org/10.1016/j.pocean.2015.05.015. Licensed under CC BY 4.0; https://creativecommons.org/licenses/by/4.0/

energy, or amplitude, of the incoming signal in a specified frequency band and to determine whether it exceeds a user-defined threshold. If the threshold within the frequency band is exceeded, the sound is scored as being present. The threshold value typically is set relative to the ambient noise in the frequency band of interest (e.g., Mellinger 2008; Ou et al. 2012). A simple energy threshold detector does not perform well when signals have low signal-to-noise ratio (SNR) or when sounds overlap. A number of techniques have been devised to overcome these problems, including spectrogram equalization (e.g., Esfahanian et al. 2017) to reduce background noise, time-varying (adaptive) detection thresholds (e.g., Morrissey et al. 2006), and using concurrent, but different, detection thresholds for different frequency bands (e.g., Brandes 2008; Ward et al. 2008). Apart from finding individual animal sounds, energy threshold detectors also have been successfully applied to the detection of animal choruses, such as those produced by spawning fish, migrating whales (Erbe et al. 2015), and chorusing insects or amphibians. These choruses are composed of many sounds from large and often distant groups of animals and so individual signals often are not detectable in them. Choruses can last for hours and significantly raise ambient levels in a species-specific frequency band (Fig. 8.6).

### 8.3.2 Spectrogram Cross-Correlation

Spectrogram cross-correlation is a well-known technique to detect sounds produced by many species, such as rockfish (genus *Sebastes*; Širović et al. 2009), African elephants (*Loxodonta africana*; Venter and Hanekom 2010), maned wolves (*Chrysocyon brachyurus*; Rocha et al. 2015), minke whales (Oswald et al. 2011), and sei whales (*Balaenoptera borealis*; Baumgartner and Fratantoni 2008). In this method, spectrograms of reference sounds from the species of interest are converted into reference coefficients, or kernels, with one kernel for each sound type (Fig. 8.7). These reference kernels then are cross-correlated with the incoming spectrogram on a frame-by-frame basis. Kernels can be a statistical representation of spectrograms of known sound types, or they can be created empirically by trial-and-error from previously analyzed recordings.

Proper selection of reference signals is critical to the performance of the detector and thus this method is only suited for detection of stereotypical sounds. Seasonal and annual variability in call structure can significantly impact performance of these detectors and so an analysis of the variability in call structure is vital when applying spectrogram cross-correlation to detect calls in long-term datasets (Širović 2016). Another

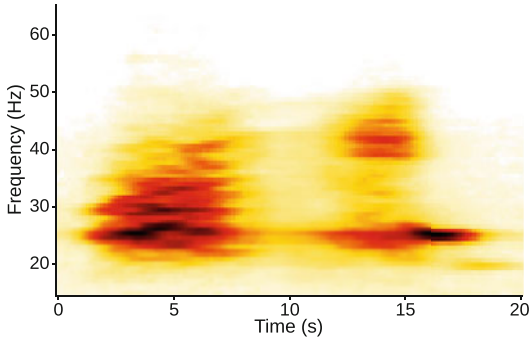



**Fig. 8.7** Spectrogram of the kernel for Omura's whales' (*Balaenoptera omurai*) doublet calls, computed as the average of over 800 hand-picked calls (Madhusudhana et al. 2020)

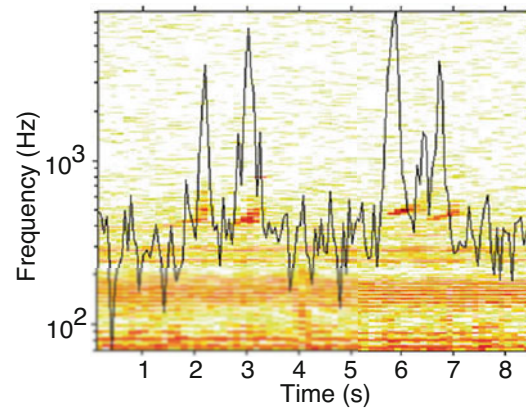



**Fig. 8.8** Spectrogram of marine mammal tonal sounds with negative entropy (black curve) overlain. Negative entropy is high when the power spectral density is concentrated in a few narrow frequency bands (Erbe and King 2008)

drawback to this method is that it can be prohibitively processor-intensive. To speed up the calculations, Harland (2008) first employed an energy threshold detector (as described above) to detect times of potential signal presence and then used spectrogram cross-correlation to detect individual signals within the flagged time periods.

### 8.3.3 Matched Filter

The matched filter approach for sound classification is similar to spectrogram cross-correlation but is performed in the time-domain. This means that the waveforms (i.e., sound pressure levels as a function of time) are correlated instead of the spectrogram. A kernel of the waveform of the sound to be detected is produced, often empirically using a high-quality recording, and then cross-correlated with the incoming signal (i.e., the time series of sound pressure). Matched filters are efficient at detecting signals in Gaussian noise (white noise), but colored noise (typical in many natural environments) poses more of a problem. As with spectrogram cross-correlation, the selection of kernels is critical to the performance of the detector. Matched filters are only appropriate for detection of well-known, stereotyped acoustic features, such as sounds produced by cane toads (Dang et al. 2008), blue whales (Stafford et al. 1998; Bouffaut et al. 2018), and beaked whales (Hamilton and Cleary 2010). Their performance suffers in the presence of even a small amount of sound variation compared to the kernel.

### 8.3.4 Spectral Entropy Detector

In general, entropy measures the disorder or uncertainty of a system. Applied to communication theory, the information entropy (also called Shannon entropy; Shannon and Weaver 1998) measures the amount of information contained in a data stream. Entropy is computed as the negative product of a probability distribution and its logarithm. Therefore, a strongly peaked probability distribution has low entropy, while a broad probability distribution has high entropy. If applied to an acoustic power spectral density distribution, entropy measures the peakedness of the power spectra and detects narrowband signals in broadband noise (Fig. 8.8). Spectral entropy has successfully been applied to animal sounds; for example, from birds, beluga whales (*Delphinapterus leucas*), bowhead whales, and walruses (Erbe and King 2008; Mellinger and Bradbury 2007; Valente et al. 2007).

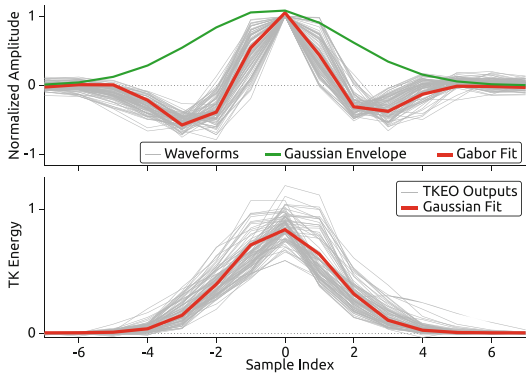



**Fig. 8.9** Waveforms of odontocete clicks and their Gabor fit (top) and TKEO outputs and Gaussian fit (bottom) (Madhusudhana et al. 2015)

### 8.3.5 Teager–Kaiser Energy Operator

The Teager–Kaiser energy operator (TKEO) is a nonlinear operator that tracks the energy of a data stream (Fig. 8.9). Operating on a time series, at any one instance, the TKEO computes the square of the sample and subtracts the product of the previous and next sample. The output is therefore high for very brief signals. The TKEO has successfully been applied to the detection of clicks, such as bat or odontocete biosonar sounds (Kandia and Stylianou 2006; Klinck and Mellinger 2011). Many biosonar signals are of Gabor type (i.e., a sinusoid modulated by a Gaussian envelope). The TKEO output of the signals is a simple Gaussian, which can be detected with simple tools, such as energy threshold detection or matched filtering (Madhusudhana et al. 2015).

### 8.3.6 Evaluating the Performance of Automated Detectors

Automated detectors can produce two types of errors: missed detections (i.e., missing a sound that exists) and false alarms (i.e., incorrectly reporting a sound that does not exist or reporting a sound that is not the target signal). There is an inevitable trade-off when choosing the acceptable rate of each. Most detectors allow the user to adjust a threshold, and depending on where this threshold is set, the probability of one type of error increases while the other decreases. The acceptability of either type of error is determined by the particular application of the detector. For example, for rare animals in critical habitats, detecting every sound, even those that are very faint, is desired. In this situation, a low threshold can be chosen that minimizes the number of missed detections; however, this can result in many false alarms. Quantification of these two errors is a useful way to evaluate the performance of an automated detector.

| | | Detector Input | |
|---|---|---|---|
| | | Signal Present | Signal Absent |
| Reported Output | Signal Present | True Positive (TP) Correct Detection | False Positive (FP) False Alarm |
| | Signal Absent | False Negative (FN) Missed Detection | True Negative (TN) Correct Rejection |

**Fig. 8.10** Confusion matrix showing the possible outcomes of a detector when a signal is present versus absent

#### 8.3.6.1 Confusion Matrices

One of the simplest and most common methods for conveying the performance of a detector (or a classifier) is a confusion matrix (i.e., a type of contingency table). A confusion matrix (Fig. 8.10) gives the number of true positives (i.e., correctly classified sounds, also called correct detections), false positives (i.e., false alarms), true negatives (i.e., correct rejections), and false negatives (i.e., missed detections).

#### 8.3.6.2 Receiver Operating Characteristic (ROC) Curve

The performance of detectors can be visualized using the receiver operating characteristic (ROC) curve. A ROC curve is a graph that depicts the trade-offs between true positives and false positives (Egan 1975; Swets et al. 2000). The false positive rate (i.e., FP/(FP+TN)) is plotted on the x-axis, while the true positive rate (i.e., TP/(TP +FN)) is plotted on the y-axis (Fig. 8.11). A curve is generated by plotting these values for the detector at different threshold values. The (0l1) point on the graph represents perfect performance: 100% true positives and no false positives.

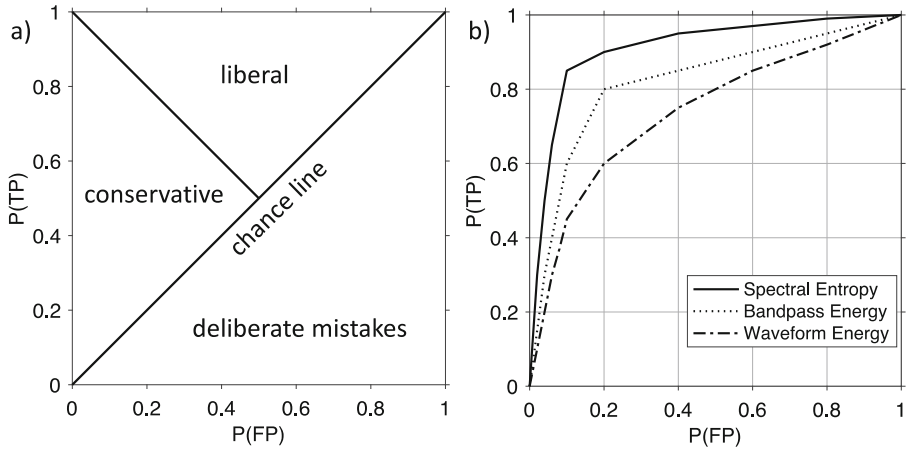



**Fig. 8.11** (**a**) Generalized receiver operating characteristic (ROC) plot, in which the probability of true positives is plotted against the probability of false positives. Areas in this graph that correspond to a liberal bias, conservative bias, and deliberate mistakes are indicated. (**b**) Example ROC curves computed during the development of automated detectors for marine mammal calls in the Arctic. The spectral entropy detector outperformed others (Erbe and King 2008)

The major diagonal in Fig. 8.11a represents performance at chance, where the probabilities of TP and FP are equal. Responses falling below the line would indicate deliberate mistakes. The minor diagonal represents neutral bias, and splits responses into conservative versus liberal. A conservative response strategy yields decreased correct detection and false alarm probabilities; a liberal response strategy yields increased correct detection and false alarm probabilities. An example ROC curve is given in Fig. 8.11b, comparing the performances of three detectors (operating on underwater acoustic recordings from the Arctic and trying to detect marine mammal calls) based on: (1) spectral entropy, (2) bandpassed energy, and (3) waveform (i.e., broadband) energy. The performance of the entropy detector surpassed that of the other two.

#### 8.3.6.3 Precision and Recall

The performance of a detector can be overestimated using a ROC curve when there is a large difference between the numbers of TPs and TNs. In addition, estimation of the number of TNs requires discrete sampling units. The duration of the discrete sampling units is often somewhat arbitrary and can lead to unrealistic differences between the numbers of TPs and TNs. In these situations, precision and recall (P-R) can provide a more accurate representation of detector performance because this representation does not rely on determining the number of true negatives (Davis and Goadrich 2006). In the P-R framework, events are scored only as TPs, FPs, and FNs.

Precision is a measure of accuracy and is the proportion of automated detections that are true detections.

$$\text{Precision} = \frac{TP}{TP + FP}$$

Recall is a measure of completeness and is the proportion of true events that are detected. This is the same as the true positive rate defined in the ROC framework.

$$\text{Recall} = \frac{TP}{TP + FN}$$

Detectors can be evaluated by plotting precision against recall (Fig. 8.12). An ideal detector would have both scores approaching a value of 1. In other words, the curve would approach the upper right-hand corner of the graph (Davis and Goadrich 2006). Precision and recall also can be

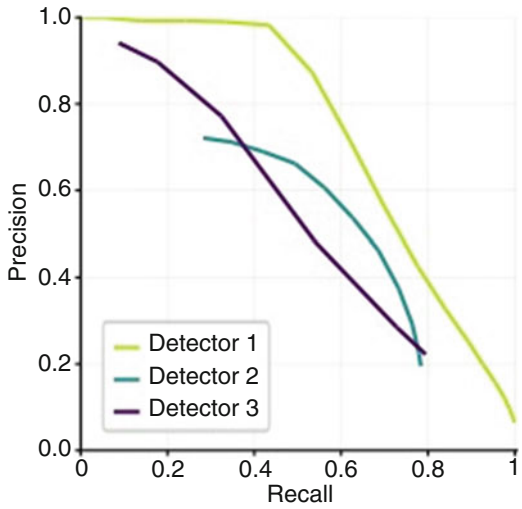



**Fig. 8.12** Precision-Recall curves for three types of detectors: (1) spectrogram cross-correlation, (2) blob detection, and (3) spectral entropy for Omura's whale calls (Madhusudhana et al. 2020)

represented by an *F*-score, which is the geometric mean of these values. The F-score can be weighted to emphasize either precision or recall when optimizing detector performance (Jacobson et al. 2013).

## 8.4 Quantitative Classification of Animal Sounds

Quantitative classification of animal sounds is based on measured features of sounds, no matter whether these are used to manually or automatically group sounds with the aid of software algorithms. These features can be measured from different representations of sounds, such as waveforms, power spectra, spectrograms, and others. A large variety of classification methods have been applied to animal sounds, many drawing from human speech analysis.

### 8.4.1 Feature Selection

The acoustic features selected and the consistency with which the measurements are taken have a significant influence on the success (or failure) of a classification algorithm. Feature sets (also called feature vectors) should provide as much information as sensible about the sounds. With today's software tools and computing power, a limitless number of features can easily be measured that would allow distinction between sounds even of the same type. Such over-parameterization can make it difficult to group like sounds, which can be just as important as distinguishing between different sounds. The challenge is to find the trade-off and produce a set of representative features for each sound type. Once the features have been selected, automating the extraction and subsequent analysis of these features reduces the time required to analyze large datasets. Some commonly used feature vectors are described below.

#### 8.4.1.1 Spectrographic Features

Perhaps the most commonly used feature vectors are those consisting of values measured from spectrograms. These measurements include, but are not limited to, frequency variables (e.g., frequency at the beginning of the sound, frequency at the end of the sound, minimum frequency, maximum frequency, frequency of peak energy, bandwidth, and presence/absence of harmonics or sidebands; Fig. 8.13; also see Chap. 4, Sect. 4.2.3), and time variables (e.g., signal duration, phrase and song length, inter-signal intervals, and repetition rate). More complex features, such as those describing the spectrographic shape of a sound (e.g., upsweep, downsweep, chevron, U-loop, inverted U-loop, or warble), slopes, and numbers and relative positions of local extrema and inflection points (places where the contour changes from positive to negative slope or vice versa) also have been used in classification. These measurements often are taken manually from spectrographic displays (e.g., by a technician using a mouse-controlled cursor). Automated techniques for extracting spectrographic measurements can be less subjective and less time-consuming, but are sometimes not as accurate as manual methods. Examples are available in the bird literature (e.g., Tchernichovski et al. 2000), bat literature (Gannon et al. 2004; O'Farrell et al. 1999), and marine mammal

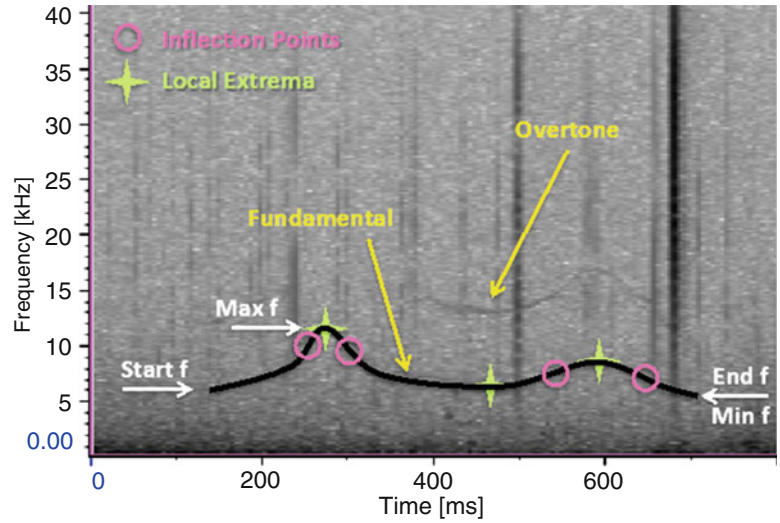



**Fig. 8.13** Spectrogram of a pilot whale (*Globicephala melas*) whistle showing the following features: Start frequency (Start f), End frequency (End f), Maximum frequency (Max f), Minimum frequency (Min f), locations of two local maxima and one local minimum in the fundamental contour, four inflection points (where the curvature changes from clockwise to counter-clockwise, or vice versa), and one overtone (Courts et al. 2020). © Courts et al.; https://www.nature.com/articles/s41598-020-74111-y/figures/5. Licensed under CC BY 4.0; https://creativecommons.org/licenses/by/4.0/

literature (e.g., Mellinger et al. 2011; Roch et al. 2011; Gillespie et al. 2013; Kershenbaum et al. 2016). Spectrographic measurements of bat calls, for example, can be extracted using *Analook* (Titley Scientific, Columbia, MO, USA), *SonoBat* (Joe Szewczak, Department of Biology, Humboldt State University, Arcata, CA, USA), or *Kaleidoscope* Pro (Wildlife Acoustics, Inc., Maynard, MA, USA), exported to an *Excel* spreadsheet (XML, CSV, and other formats), classified using machine learning algorithms, and compared to a reference library for identification.

#### 8.4.1.2 Cepstral Features

Cepstral coefficients are spectral features of bioacoustic signals commonly used in human speech processing (Davis and Mermelstein 1980). These features are based on the source-filter model of human speech analysis, which has been applied to many different animal species (Fitch 2003). Cepstral coefficients are well-suited for statistical pattern-recognition models because they tend to be uncorrelated (Clemins et al. 2005), which significantly reduces the number of parameters that must be estimated (Picone 1993). Cepstral coefficients are calculated by computing the Fourier transform in successive time windows over the recorded pressure time series of a sound (see Chap. 4). The frequency axis then is warped by multiplying the spectrum with a series of $n$ filter functions at appropriately spaced frequencies. This is done because there is evidence that many animals perceive frequencies on a logarithmic scale, in a similar fashion to humans (Clemins et al. 2005). The output of the frequency band filters is then used as input to a discrete cosine transform, which results in an $n$-dimensional cepstral feature vector (Picone 1993; Clemins et al. 2005; Roch et al. 2007, 2008).

Using cepstral feature space allows the timbre of sounds to be captured, a quality that is lost when extracting parameters from spectrograms. Roch et al. (2007) developed an automated classification system based on cepstral feature vectors extracted for whistles, burst-pulse sounds, and clicks produced by short- and long-beaked

common dolphins (*Delphinus* spp.), Pacific white-sided dolphins (*Lagenorhynchus obliquidens*), and bottlenose dolphins (*Tursiops truncatus*). The system did not rely on specific sound types and had no requirement for separating individual sounds. The system performed relatively well, with correct classification scores of 65–75%, depending on the partitioning of the training- and test-data. Cepstral feature vectors also have been used as input to classifiers for many other animal species, including groupers (*Epinephelus guttatus, E. striatus, Mycteroperca venenosa, M. bonaci;* Ibrahim et al. 2018), frogs (Gingras and Fitch 2013), song birds (Somervuo et al. 2006), African elephants (Zeppelzauer et al. 2015), and beluga, bowhead, gray (*Eschrichtius robustus*), humpback, and killer (*Orcinus orca*) whales, and walrus (Mouy et al. 2008). Cepstral features appear to be a promising alternative to the traditional time- and frequency-parameters measured from spectrograms as input to classification algorithms. However, cepstral features are relatively sensitive to the SNR, the signal's phase, and modeling order (Ghosh et al. 1992).

Noda et al. (2016) used mel-frequency cepstral coefficients and random forest analyses to classify sounds produced by 102 species of fish and compared the performance of three classifiers: k-nearest neighbors, random forest, and support vector machines (SVMs). The mel-frequency cepstrum (or cepstrogram) is a form of acoustic power spectrum (or spectrogram) that is computed as a linear cosine transform of a log-power spectrum that is presented on a nonlinear mel-scale of frequency. The mel-scale resembles the human auditory system better than the linearly-spaced frequency bands of the normal cepstrum. All three classifiers performed similarly, with average classification accuracy ranging between 93% and 95%.

### 8.4.2 Statistical Classification of Animal Sounds

For some sounds, qualitative classification is sufficient. Janik (1999) reported that humans were able to identify dolphin signature whistles more reliably than computer methods. A problem with qualitative classification of sounds in a repertoire (and taxonomy in general), however, is that some listeners are "splitters" and other listeners are "lumpers." So, even researchers on the same project could classify an animal's sound repertoire differently. One way to avoid individual researcher differences in classification is to use graphical, statistical, and computer-automated methods that objectively sort and compare measured variables that describe the sounds. A variety of statistical methods can be employed to classify animal sounds into categories (Frommolt et al. 2007). Below are brief descriptions of some of the statistical methods that are commonly used for classification of animal sounds.

#### 8.4.2.1 Parametric Clustering

Parametric cluster analysis produces a dendrogram (i.e., classification tree) that organizes similar sounds into branches of a tree. A distance matrix also is generated, which gives correlation coefficients between all variables in the dataset. The resulting distance index ranges from 0 (very similar sounds) to 1 (totally dissimilar sounds). The matrix can then be joined by rows or columns to examine relationships. The type of linkage and type of distance measurement can be selected to find the best fit for a particular dataset (Zar 2009).

Cluster analysis has been used to classify sound types in several species, including owls (Nagy and Rockwell 2012), mice (Hammerschmidt et al. 2012), rats (*Rattus norvegicus*, Takahashi et al. 2010), African elephants (Wood et al. 2005), and primates (Hammerschmidt and Fischer 1998). In a study of six populations of the neotropical frog (*Proceratophrys moratoi*) in Brazil, Forti et al. (2016) measured spectrographic variables from calls produced by males and performed cluster analysis to examine similarities in acoustic traits (based on the Bray–Curtis index of acoustic similarity) across the six locations (Fig. 8.14). Baptista and Gaunt (1997) used hierarchical cluster analysis of correlation coefficients of several acoustic parameters to categorize sounds of the sparkling violet-eared hummingbird (*Colibri*

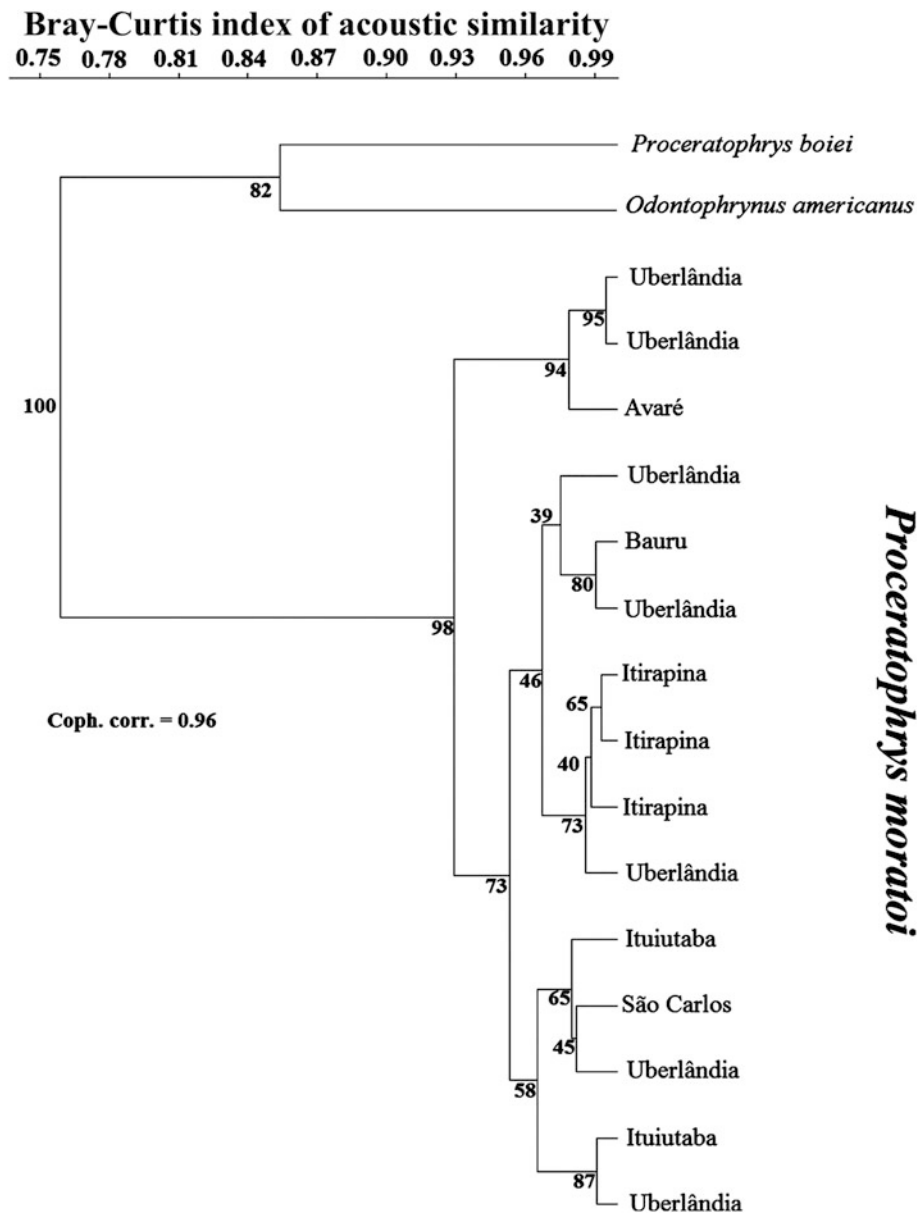



**Fig. 8.14** Dendrogram from a hierarchical cluster analysis of the call similarities between 15 male *Proceratophrys moratoi* from different sites and two other Odontophrynidae species (Forti et al. 2016). © Forti et al.; https://peerj.com/articles/2014/. Licensed under CC BY 4.0; https://creativecommons.org/licenses/by/4.0/

*coruscans*), which is found in two neighboring assemblages in their study area. A matrix of sound similarity values obtained from spectral cross-correlation of these birds' songs indicated similar sound types from the two areas. Yang et al. (2007) used cluster analysis to examine syllable sharing between individuals of Anna's hummingbird (*Calypte anna*). They identified 38 syllable types in songs of 44 males, which clustered into five basic syllable categories: "Bzz," "bzz," "chur," "ZWEE," and "dz!". Also, microgeographic song variation patterns were

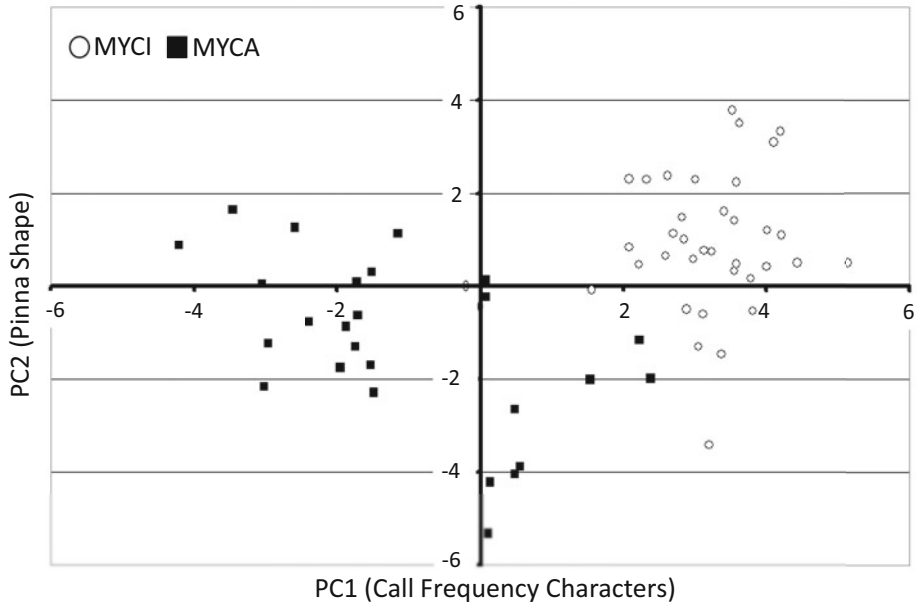



**Fig. 8.15** Plot showing the results of principal component analysis, in which two cryptic species of myotis bats (California myotis, *Myotis californicus,* MYCA, black squares; western small-footed bat, *M. ciliolabrum,* MYCI, hollow circles) were distinguished by differences in ear height and characteristic frequency of their echolocation signals. Plotted is characteristic frequency versus signal duration for these species recorded from field sites in New Mexico and Arizona, USA

found in that nearest neighbors sang more similar songs than non-neighbors. Pozzi et al. (2010) used several acoustic variables to group black lemur (*Eulemur macaco macaco*) sounds into categories, including the frequencies of the fundamental and of the first three harmonic overtones (measured at the start, middle, and end of each call), and the total duration. The agreement of this analysis with manual classification was high (>88.4%) for six of eight categories.

#### 8.4.2.2 Principal Component Analysis

Principal component analysis (PCA) is a multivariate statistical method that examines a set of measurements such as the feature vectors discussed earlier in Sect. 8.4. These features may well be correlated. For example, bandwidth is sometimes correlated with maximum frequency, or the number of inflection points can be correlated with signal duration (Ward et al. 2016). PCA performs an orthogonal transformation that converts the potentially correlated variables (i.e., the features) into a set of linearly uncorrelated variables (i.e., the principal components; Hotelling 1933; Zar 2009). The principal components are linear combinations of the original variables (features). Plotting the principal components against each other shows how the measurements cluster.

For example, by examining bat biosonar signals in multivariate space, bat species that are very similar in external appearance can be distinguished. Using PCA, Gannon et al. (2001) found ear height and characteristic frequency were correlated, along with duration of the signal (Fig. 8.15).

As another example, Briefer et al. (2015) categorized emotional states associated with variation in whinnies from 20 domestic horses (*Equus ferus*) using PCA. They designed four situations to elicit different levels of emotional arousal that were likely to stimulate whinnies: separation (negative situation) and reunion (positive situation) with either all group members (high

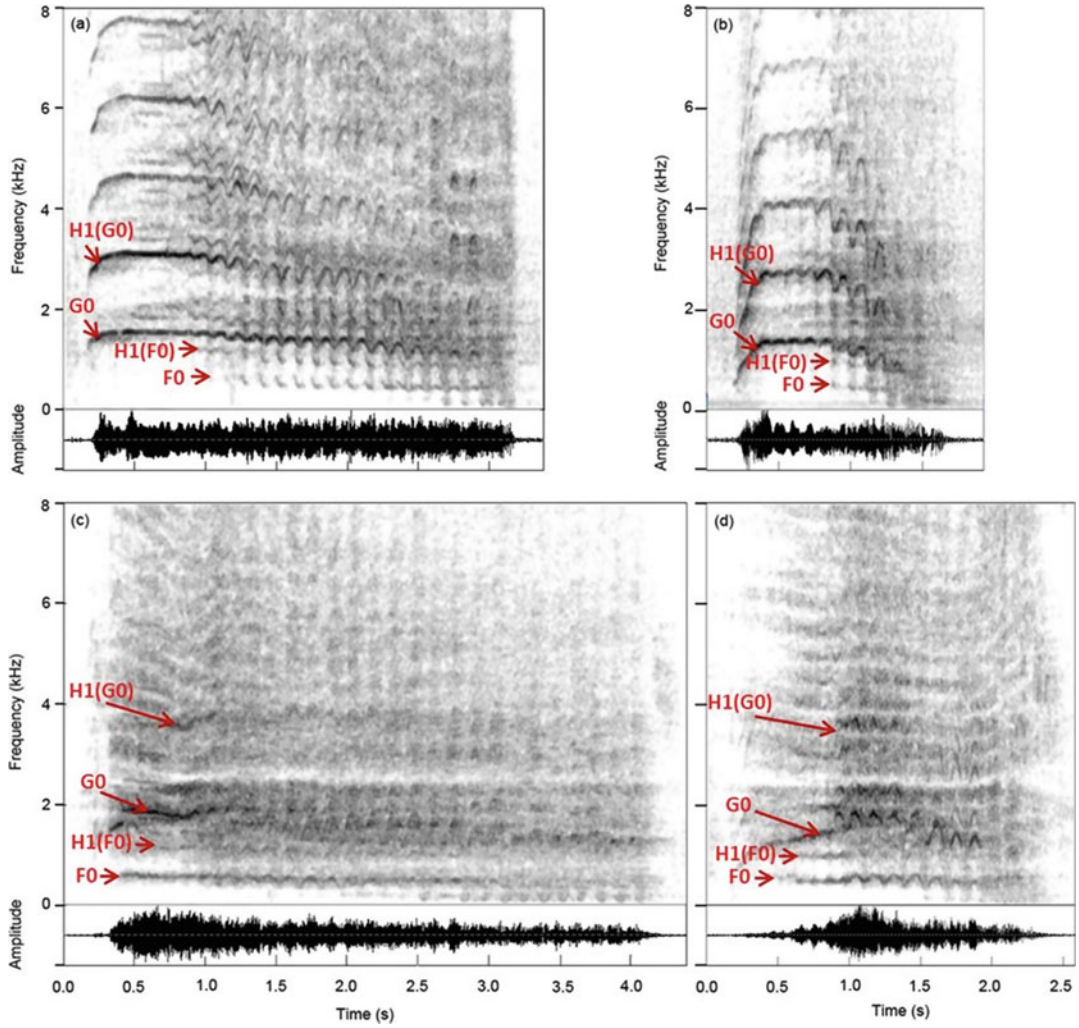



**Fig. 8.16** Spectrograms and oscillograms of horse whinnies in negative (**a**, **c**) and positive (**b**, **d**) situations emitted by two different horses. Red arrows point to fundamental frequencies (F0, G0) and first overtones (H1). Negative whinnies (**a**, **c**) are longer in duration and have higher G0 fundamentals than positive whinnies (**b**, **d** Briefer et al. 2015). © Briefer et al.; https://www.nature.com/articles/srep09989/figures/3. Licensed under CC BY 4.0; http://creativecommons.org/licenses/by/4.0/

emotional arousal) or only one group member (moderate emotional arousal). The authors measured 21 acoustic features from whinnies (Fig. 8.16). PCA transformed the feature vectors into six principal components that accounted for 83% of the variance in the original dataset.

#### 8.4.2.3 Discriminant Function Analysis

In discriminant function analysis (DFA), canonical discriminant functions are calculated using variables measured from a training dataset. One canonical discriminant function is produced for each sound type in the dataset. Variables measured from sounds in the test dataset are then substituted into each function and each sound type is classified according to the function that produced the highest value. Because DFA is a parametric technique, it is assumed that input data have a multivariate normal distribution with the same covariance matrix (Afifi and Clark 1996;

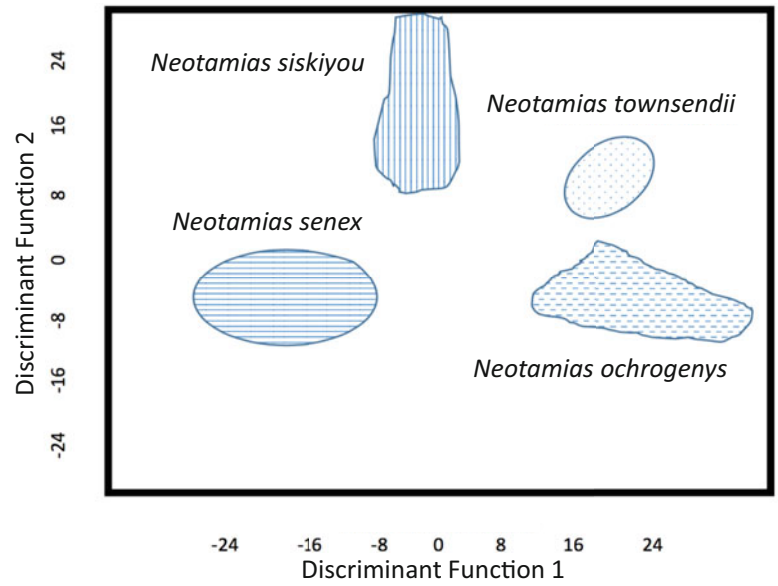



**Fig. 8.17** Plot resulting from discriminant function analysis. Four species of Townsend-group chipmunks (Townsend's chipmunk, *Neotamias townsendii*; Siskiyou chipmunk, *N. siskiyou*; Allen's chipmunk, *N. senex*; and yellow-cheeked chipmunk, *N. ochrogenys*) in northern California, USA, produced discernibly different sounds. Discriminant function 1 was dominated by differences in maximum frequency of the signal and discriminant function 2 was most influenced by temporal features including total signal length and the number of signals emitted by a chipmunk during a signaling bout

Zar 2009). Violations of these assumptions can create problems with some datasets. One of the main weaknesses of DFA for animal sound classification is that it assumes classes are linearly separable. Because a linear combination of variables takes place in this analysis, the feature space can only be separated in certain, restricted ways that are not appropriate for all animal sounds. Figure 8.17 shows the DFA separation of California chipmunk (genus *Neotamias*) taxa that are morphologically similar but acoustically different, using six variables measured from their sounds.

#### 8.4.2.4 Classification Trees

Classification tree analysis is a non-parametric statistical technique that recursively partitions data into groups known as "nodes" through a series of binary splits of the dataset (Clark and Pregibon 1992; Breiman et al. 1984). Each split is based on a value for a single variable and the criteria for making splits are known as primary splitting rules. The goal for each split is to divide the data into two nodes, each as homogeneous as possible. As the tree is grown, results are split into successively purer nodes. This continues until each node contains perfectly homogeneous data (Gillespie and Caillat 2008). Once this maximal tree has been generated, it is pruned by removing nodes and examining the error rates of these smaller trees. The smallest tree with the highest predictive accuracy is the optimal tree (Oswald et al. 2003).

Tree-based analysis provides several advantages over some of the other classification techniques. It is a non-parametric technique; therefore, data do not need to be normally distributed as required for other methods, such as DFA. In addition, tree-based analysis is a simple and naturally intuitive way for humans to classify sounds. It is essentially a series of true/false questions, which makes the classification process transparent. This allows easy examination of which variables are most important in the classification process. Tree-based analysis also

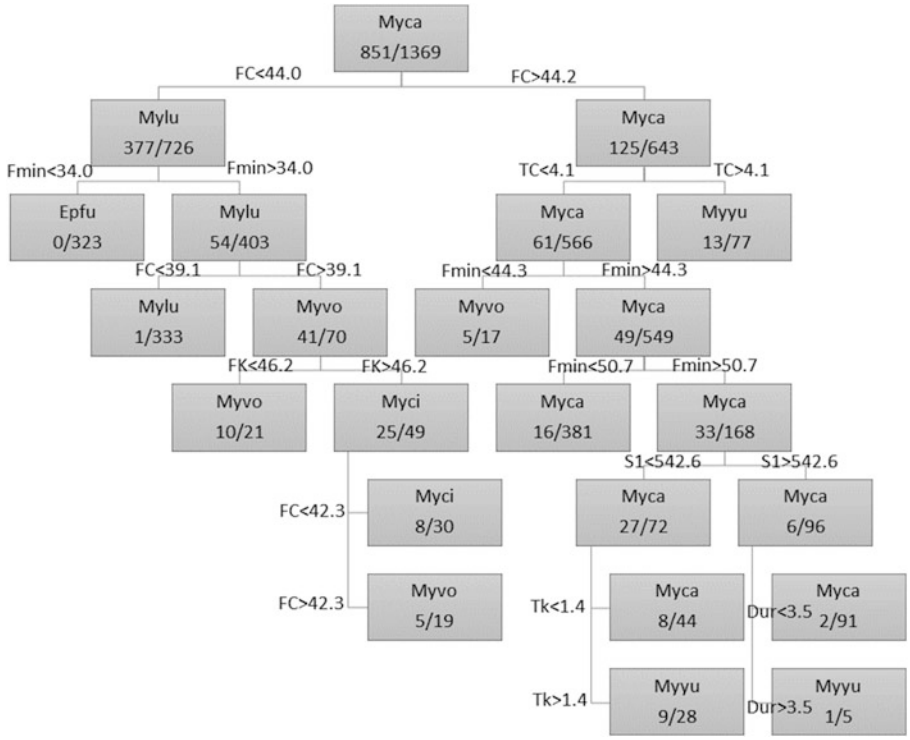



**Fig. 8.18** Classification tree grown using *Splus* computer software (version S-PLUS 6.2 2003, TIBCO Software Inc., Palo Alto, CA, USA) from 1369 bat calls. The pruned tree used variables measured from each bat call: duration (DUR), minimum frequency (Fmin), characteristic frequency (Fc; i.e., frequency at the flattest part of the call), frequency at the "knee" of the call (Fk), time of Fc, time at Fk, and slope (S1). Along the tangents between boxes are values for variables used to split the nodes (for instance, Fmin is minimum frequency). The fraction below each box is the misclassification rate (e.g., 1/5 = 20% misclassification rate). The tree has 12 terminal nodes defining the branches, resulting in a classification designation for each species (Gannon et al. 2004)

accommodates for a high degree of diversity within classes. For example, if a species produces two or more distinct sound types, a tree-based analysis can create two different nodes. In other classification techniques, different sound types within a species simply act to increase variability and make classification more difficult. Finally, surrogate splitters are provided at each node (Oswald et al. 2003). Surrogate splitters closely follow primary splitting rules and can be used in cases when the primary splitting variable is missing. Therefore, sounds can be classified even if data for some variables are missing due to noise or other factors.

To address some controversy as to whether closely related species of myotis bats could be differentiated by their sounds, Gannon et al. (2004) completed an analysis of echolocation pulses from free-flying, wild bats. Fig. 8.18 is a classification tree grown from nearly 1400 calls using at least seven variables measured from each call. The tree produced terminal nodes identified to species (MYVO is *Myotis volans*, MYCA *M. californicus,* etc.). In this study, recordings were made under field conditions where sounds were affected by the environment, Doppler shift, and diversity of equipment. Still, classification trees worked well to predict group membership and additional techniques, such as DFA, were able to distinguish five *Myotis* species acoustically with greater than 75% accuracy (greater than 90% in most instances).

Classification trees have been applied to marine mammal sounds by several researchers,

with promising results. Fristrup and Watkins (1993) used tree-based analysis to classify the sounds of 53 species of marine mammal (including mysticetes, odontocetes, pinnipeds, and manatees). Their correct classification score of 66% was 16% higher than the score obtained when applying DFA to the same dataset. The whistles of nine delphinid species were correctly classified 53% of the time by Oswald et al. (2003) using tree-based analysis. Oswald et al. (2007) subsequently applied classification tree analysis to the whistles of seven species and one genus of marine mammal, resulting in a correct classification score of 41%. This score was improved slightly, to 46%, when classification decisions were based on a combination of classification tree and DFA results. Gannier et al. (2010) used classification trees to identify the whistles of five delphinid species recorded in the Mediterranean, with a correct classification score of 63%. Finally, Gillespie and Caillat (2008) classified the clicks of Blainville's beaked whales (*Mesoplodon densirostris)*, short-finned pilot whales (*Globicephala macrorhynchus*), and Risso's dolphins (*Grampus griseus*). Their tree-based analysis classified 80% of clicks to the correct species.

#### 8.4.2.5 Nonlinear Dimensionality Reduction

Clustering techniques described above require that certain features or measurements, as appropriate for the problem domain, be available beforehand. They are gathered from sound recordings either manually (e.g., number of inflection points in whistle contours, number of harmonics) or using signal processing tools (e.g., peak frequency, energy), or both. Manual extraction of features is usually time-consuming and often inefficient, especially when dealing with recordings covering large spatial and temporal scales. Automated extraction of measurements improves efficiency and eliminates the risk of human biases. However, when recordings contain a lot of confounding sounds or have extreme noise variations, reliability and accuracy of the measurements can become questionable and can have adverse effects on clustering outcomes. Regardless of whether manual or automated approaches were employed, the resulting limited set of chosen features or measurements are essentially representations of the underlying data in a reduced space. Such dimensionality reduction is typically aimed at making the downstream task of clustering (with PCA, DFA, etc.) computationally tractable.

In recent years, nonlinear dimensionality reduction methods have gained widespread popularity, specifically in applications for exploring and visualizing very high-dimensional data. Originally popular for processing image-like data in the field of machine learning, these methods bring about dimensionality reduction without requiring one to explicitly choose and extract features. The methods can be easily adapted for processing bioacoustic recordings wherein the qualitative cluster structure (i.e., similarities in the visually identifiable information) in spectrogram-like data (e.g., mel-spectrogram or cepstrogram) containing hundreds or thousands of time-frequency points is effectively captured in an equivalent 2- or 3-dimensional space (e.g., Sainburg et al. 2019; Kollmorgen et al. 2020).

One of the earlier methods for capturing nonlinear structure, the t-distributed stochastic neighbor embedding (t-SNE; van der Maaten and Hinton 2008) is based on non-convex optimization. It computes a similarity measure between pairs of points (data samples) in the original high-dimensional space and in the reduced space, then minimizes the Kullback–Leibler divergence between the two sets of similarity measures. t-SNE tries to preserve distances in a neighborhood whereby points close together in the high-dimensional space have a high probability of staying close in the reduced space. The *Bird Sounds* project (Tan and McDonald 2017) presents an excellent demonstration of using t-SNE for organizing thousands of bird sound spectrograms in a 2-dimensional similarity grid.

Some of the shortcomings of t-SNE were addressed in a newer method called uniform manifold approximation and projection (UMAP; McInnes et al. 2018). UMAP is backed with a strong theoretical framework. While effectively capturing local structures like t-SNE, UMAP

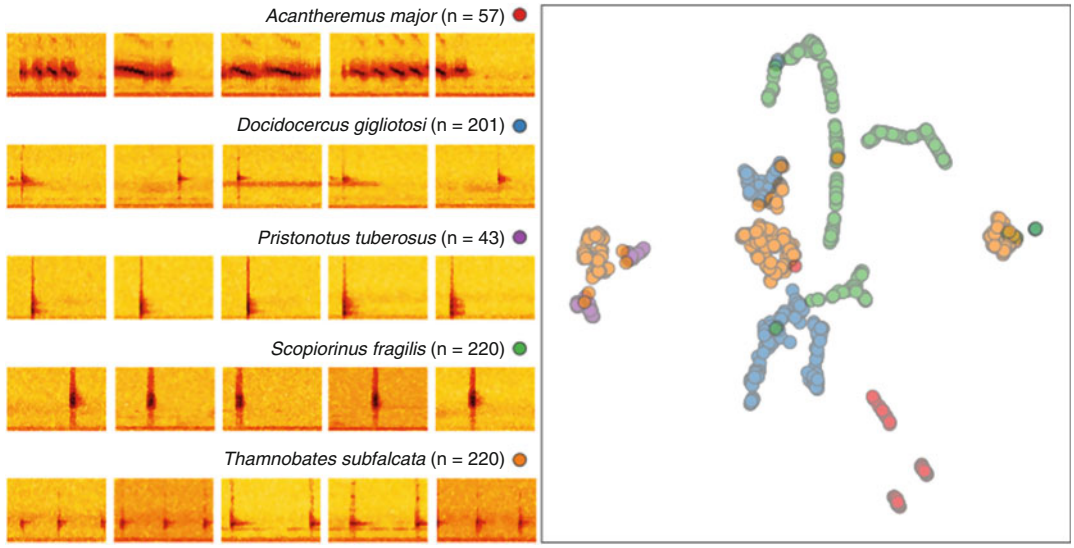



**Fig. 8.19** Demonstration of clustering katydid sounds using UMAP. Randomly chosen samples of call spectrograms of the five species considered are shown on the left, and clustering outcomes are shown on the right. The clustering activity has successfully captured both inter-species and intra-species variations

also offers a better promise for preserving global structures (inter-cluster relationships). UMAP processes data faster and is capable of handling very large dimensional data. Fig. 8.19 is a demonstration of the use of UMAP for clustering sounds of five species of katydids (Tettigoniidae) from Panamanian rainforest recordings (Madhusudhana et al. 2019). Inputs to UMAP clustering comprised of spectrograms (dimensions 216h x 469w) computed from 1-s clips containing katydid call(s). The inputs often contained confounding sounds and varying noise levels. The clustering results, however, demonstrate the utility of UMAP as a quick means to effective clustering. UMAP has also been used, in combination with a pre-trained neural network, for assessing habitat quality and biodiversity variations from soundscape recordings across different ecosystems (Sethi et al. 2020).

We have presented here two popular methods that are currently trending in this field of research. There are, however, other alternatives available including earlier methods such as isomap (Tenenbaum et al. 2000) and diffusion map (Coifman et al. 2005), newer variants of t-SNE (e.g., Maaten 2014; Linderman et al. 2017), and some modern variants of variational autoencoders (Kingma and Welling 2013).

## 8.4.3 Model Based Classification

### 8.4.3.1 Artificial Neural Networks

Artificial neural networks (ANNs) were developed by modeling biological systems of information-processing (Rosenblatt 1958) and became very popular in the areas of word recognition in human speech studies (e.g., Waibel et al. 1989; Gemello and Mana 1991) and character or image-recognition (e.g., Fukushima and Wake 1990; Van Allen et al. 1990; Belliustin et al. 1991) in the 1980s. Since that time, ANNs have been used successfully to classify a number of complex signal types, including quail crows (*Coturnix* spp., Deregnaucourt et al. 2001), alarm sounds of Gunnison's prairie dogs (*Cynomys gunnisoni*, Placer and Slobodchikoff 2000), stress sounds by domestic pigs (*Sus scrofa domesticus*, Schon et al. 2001), and dolphin echolocation clicks (Roitblat et al. 1989; Au and Nachtigall 1995).

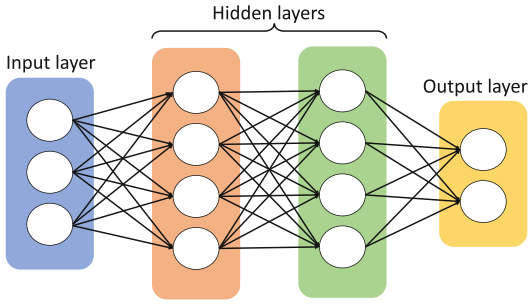



**Fig. 8.20** Diagram of the structure of an artificial neural network

In their primitive forms, there are 20 or more basic architectures of ANNs (see Lippman 1989 for a review). Each ANN approach results in trade-offs in computer memory and computation requirements, training complexity, and time and ease of implementation and adaptation (Lippman 1989). The choice of ANN depends on the type of problem to be solved, size and complexity of the dataset, and the computational resources available. All ANNs are composed of units called neurons and connections among them. They typically consist of three or more neuron layers: one input layer, one output layer, and one or more hidden layers (Fig. 8.20). The input layer consists of $n$ neurons that code for $n$ features in the feature vector representing the signal ($X_1 \ldots X_n$). The output layer consists of $k$ neurons representing the $k$ classes. The number of hidden layers between the input and output layers, as well as the number of neurons per layer, is empirically chosen by the researcher. Each connection among neurons in the network is associated with a weight-value, which is modified by successive iterations during the training of the network.

ANNs are promising for automatic signal classification for several reasons. First, the input to an ANN can range from feature vectors of measurements taken from spectrograms or waveforms, to frequency contours, to complete spectrograms. Second, ANNs serve as adaptive classifiers which learn through examples. As a result, it is not necessary to develop a good mathematical model for the underlying signal characteristics before analysis begins (Ghosh et al. 1992). In addition, ANNs are nonlinear estimators that are well-suited for problems involving arbitrary distributions and noisy input (Ghosh et al. 1992; Potter et al. 1994).

Dawson et al. (2006) used artificial neural networks as a means to classify the chick-a-dee-dee-dee call of the black-capped chickadee *(Poecile atricapillus)*, which contains four note types carrying important functional roles in this species. In their study, an ANN first was trained to identify the note type based on several acoustic variables and then correctly classified recordings of the notes with 98% accuracy. The performance of the network was compared with classification using DFA, which also achieved a high level of correct classification (95%). The authors concluded that "there is little reason to prefer one technique over another. Either method would perform extremely well as a note-classification tool in a research laboratory" (Dawson et al. 2006).

Placer and Slobodchikoff (2000) used artificial neural networks to classify alarm sounds of Gunnison's prairie dogs (*Cynomys gunnisoni*) to predator species with a classification accuracy of 78.6 to 96.3%. The ANN identified unique signals for four different species of predators: red-tailed hawk (*Buteo jamaicensis*), domestic dog (*Canis familiaris*), coyote (*Canis latrans*), and humans (*Homo sapiens*).

Deecke et al. (1999) used artificial neural networks to examine dialects in underwater sounds of killer whale pods. The neural network extracted the frequency contours of one sound type shared by nine social groups of killer whales and created a neural network similarity index. Results were compared to the sound similarity judged by three humans in pair-wise classification tasks. Similarity ratings of the neural network mostly agreed with those of the humans, and were significantly correlated with the killer whale group, indicating that the similarity indices were biologically meaningful. According to the authors, "an index based on neural network analysis therefore represents an objective and repeatable means of measuring acoustic similarity, and allows comparison of results across studies, species, and time" (Deecke et al. 1999).

The greater potential of ANNs remained largely untapped for many years, in part due to prevailing limitations in computational capabilities. In the mid-1980s, backpropagation paved a way for efficiently training multi-layer ANNs (Rumelhart et al. 1986). Backpropagation, an algorithm for supervised learning of the weights in an ANN using gradient descent, greatly facilitated development of deeper networks (having many hidden layers). Many classes of deep neural networks (DNNs; LeCun et al. 2015) such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs) became easier to train. While the aforementioned ANN approaches often require hand-picked features or measurements as inputs, DNNs trained with backpropagation demonstrated the ability to learn good internal representations from raw data (i.e., the hidden layers captured non-trivial representations effectively). In their landmark work on using CNNs for the automatic recognition of handwritten digits, LeCun et al. (1989a, b) used backpropagation to learn convolutional kernel coefficients directly from images. Over the past two decades, advances in computing technology, especially the wider availability of graphics processing units (GPUs), have considerably accelerated machine learning (ML) research in many disciplines such as computer vision, speech processing, natural language processing, recommendation systems, etc. Shift invariance is an attractive characteristic of CNNs, which makes them suitable for analyzing visual imagery (LeCun et al. 1989a, b, 1998). CNN-based solutions have consistently dominated many of the large-scale visual recognition challenges. As such, several competing architectures of CNNs have been developed: AlexNet (Krizhevsky et al. 2017), ResNet (He et al. 2016), DenseNet (Huang et al. 2017), etc. Some of these architectures have become the state-of-the-art in computer vision applications such as face recognition, emotion detection, object extraction, scene classification, and also in conservation applications (e.g., species identification in camera trap data, land-use monitoring in aerial surveys). Given the image-like nature of time-frequency representations of acoustic signals (e.g., spectrogram), many of the successes of CNNs in computer vision have been replicated in the field of animal bioacoustics. In contrast to CNNs, RNNs are better suited for processing sequence inputs. RNNs contain internal states (memory) that allow them to "learn" temporal patterns. However, their utility is limited by the "vanishing gradient problem," wherein the gradients (from the gradient descent algorithm) of the network's output with respect to the weights in the early layers become extremely small. The problem is overcome in modern flavors of RNNs such as long short-term memory (LSTM; Hochreiter and Schmidhuber 1997) networks and gated recurrent unit (GRU; Cho et al. 2014) networks.

These types of ML solutions are heavily data-driven and often require large quantities of training samples. Typically, the training samples are time-frequency representations (e.g., spectrogram or mel-spectrogram) of short clips of recordings (e.g., Stowell et al. 2016; Shiu et al. 2020). Robustness of the resulting models are improved by ensuring that the inputs adequately cover possible variations of the target signals and of the ambient background conditions. Data scientists employ a variety of data augmentation techniques to overcome data shortage. Some examples include introducing synthetic variations such as infusion of Gaussian noise, shifting in time (horizontal shift) and frequency content (vertical shift) (Jaitly and Hinton 2013; Ko et al. 2015; Park et al. 2019). The training process, which involves successively lowering a loss function iteratively using the backpropagation algorithm, is usually computationally intensive and is often sped up with the use of GPUs.

DNNs have been used in the automatic recognition vocalizations of insects (e.g., Madhusudhana et al. 2019), fish (e.g., Malfante et al. 2018), birds (e.g., Stowell et al. 2016; Goëau et al. 2016), bats (e.g., Mac Aodha et al. 2018), marsupials (e.g., Himawan et al. 2018), primates (e.g., Zhang et al. 2018), and marine mammals (e.g., Bergler et al. 2019). CNNs have been used in the recognition of social calls, song calls, and whistles (e.g., Jiang et al. 2019; Thomas et al. 2019). While typical 2-dimensional CNNs have

been successfully used in the detection of echolocation clicks (e.g., Bermant et al. 2019), 1-dimensional CNNs (with waveforms as inputs) have been attempted as well (e.g., Luo et al. 2019). CNNs and LSTM networks have been compared in an application for classifying grouper species (Ibrahim et al. 2018) where the authors observed similar performances between the two models. Shiu et al. (2020) attempted combining a CNN with a GRU network for detecting North Atlantic right whale (*Eubalaena glacialis*) calls. Madhusudhana et al. (2021) incorporated long-term temporal context by combining independently trained CNNs and LSTM networks and achieved notable improvements in recognition performance. An attractive approach for developing recognition models is the use of transfer learning technique (Torrey and Shavlik 2010), where components of an already trained model are reused. Typically, weights of the early layers of a pre-trained network are frozen (no longer trainable) and the model is adapted to the target domain by training only the leaf nodes with data from the target domain. Zhong et al. (2020) used transfer learning to produce a CNN model for classifying the calls of a few species of frogs and birds.

#### 8.4.3.2 Random Forest Analysis

A random forest is a collection of many (hundreds or thousands) individual classification trees, which are grown without pruning. Each tree is different from every other tree in the forest because at each node, the variable to be used as a splitter is chosen from a random subset of the variables (Breiman 2001). Each tree in the forest produces a predicted category for the sound to be classified as, and the sound is ultimately classified as the category that was predicted by the majority of trees. Random forests are often more accurate than single classification trees because they are robust to over-fitting and stable to small perturbations in the data, correlations between predictor variables, and noisy predictor variables. Random forests perform well on polymorphic categories such as the variety of flight calls produced by many bird species (e.g., Liaw and Wiener 2002; Cutler et al. 2007; Armitage and Ober 2010; Ross and Allen 2014).

One of the advantages of a random forest analysis is that it provides information on the degree to which each one of the input variables contributes to the final species classification. This information is given by the Gini index and is known as the Gini variable importance. The Gini index is calculated based on the "purity" of each node in each of the classification trees, where purity is a measure of the number of whistles from different species in a given node (Breiman et al. 1984). Smaller Gini indices represent higher purity. When a random forest analysis is run, the algorithm assigns splitting variables so that the Gini index is minimized at each node (Oh et al. 2003). When a forest has been grown, the Gini importance value is calculated for each variable by summing the decreases in Gini index from one node to the next each time the variable is used. Variables are ranked according to their Gini importance values—those with the highest values contribute the most to the random forest model predictions. Random forests also produce a proximity measure, which is the fraction of trees in which particular observations end up in the same terminal nodes. This measure provides information about the similarity of individual observations because similar observations should end up in the same terminal nodes more often than dissimilar observations (Liaw and Wiener 2002).

Armitage and Ober (2010) compared the classification performance of random forests, support vector machines (SVMs), artificial neural networks, and DFA for bat echolocation signals and found that, with the exception of DFA, which had the lowest classification accuracy, all classifiers performed similarly. Keen et al. (2014) compared the performance of four classification algorithms using spectrographic measurements (spectrographic cross-correlation, dynamic time-warping, Euclidean distance, and random forest) for flight calls from four warbler species. In this study, random forests produced the most accurate results, correctly classifying 68% of calls.

Oswald et al. (2013) compared classifiers generated using DFA versus random forest classifiers for whistles produced by eight delphinid species recorded in the tropical Pacific Ocean and found that random forests resulted in the highest overall correct classification score. Rankin et al. (2016) trained a random forest classifier for five delphinid species in the California Current ecosystem. This classifier used information from whistles, clicks, and burst-pulse sounds and correctly classified 84% of acoustic encounters. Both Oswald et al. (2013) and Rankin et al. (2016) used spectrographic measurements as input variables for their classifiers.

#### 8.4.3.3 Gaussian Mixture Models

Gaussian Mixture Models (GMMs) are used commonly to model arbitrary distributions as linear combinations of parametric variables. They are appropriate for species identification when there are no expectations, such as the sequence of sounds (Roch et al. 2007). To create a GMM, a set of $n$ normal distributions with separate means and diagonal covariance matrices are scaled by weight-factors $c_i$ ($1 < i < n$). The sum over all $c_i$ must be 1 to ensure that the GMM represents a probability distribution (Huang et al. 2001; Roch et al. 2007, 2008). The number of mixtures in the GMM is chosen empirically and its parameters are estimated using an iterative algorithm, such as the Expectation Maximization algorithm (Moon 1996). Once a GMM has been trained, likelihood is computed for each sound type and a log-likelihood-ratio test is used to decide the species (Roch et al. 2008).

Gingras and Fitch (2013) used GMMs to classify male advertisement songs of four genera of anurans (*Bufo*, *Hyla*, *Leptodactylus, Rana*) based on spectral features and mel-frequency cepstral coefficients. The GMM based on spectral features resulted in 60% true positives and 13% false positives, and the GMM based on mel-frequency cepstral coefficients resulted in 41% true positives and 20% false positives. Somervuo et al. (2006) correctly classified 55–71% of song fragments from 14 different species of birds based on mel-frequency cepstral coefficients. The correct classification score depended on the number of cepstral coefficients and the number of Gaussian mixtures in the model. Lee et al. (2013) used GMMs to classify song segments of 28 species of birds based on image-shape features instead of traditional spectrographic features. This approach resulted in 86% or 95% classification accuracy for 3- or 5-s birdsong segments, respectively.

Roch et al. (2008) classified clicks produced by Blainville's beaked whales, pilot whales, and Risso's dolphins using a GMM. Correct classification scores for these three species were 96.7%, 83.2%, and 99.9%, respectively. Brown and Smaragdis (2008, 2009) used GMMs to classify sounds of killer whales, resulting in up to 92% agreement with 75 perceptually created categories of sound types, depending on the number of cepstral coefficients and Gaussians in the estimate of the probability density function. GMMs were used to classify the A and B type sounds produced by blue whales in the Northeast Pacific (McLaughlin et al. 2008), and six marine mammal species (Mouy et al. 2008) recorded in the Chukchi Sea: bowhead whales, humpback whales, gray whales, beluga whales, killer whales, and walruses. Both studies reported that their classifiers worked very well, but correct classification scores were not provided.

#### 8.4.3.4 Support Vector Machines

Support vector machines (SVMs) are a rich family of learning algorithms based on Vapnik's (1998) statistical learning theory. An SVM works by mapping features measured from sounds into a high-dimensional feature space. The SVM then finds the optimal hyperplane (function) that maximizes the separation among classes with the lowest number of parameters and the lowest risk of error. This approach attempts to meet the goal of minimizing both the training error and the complexity of the classifier (Mazhar et al. 2007). The best hyperplane is one that maximizes the distance between the hyperplane and the nearest data points belonging to different classes. The support vectors are the data points that determine the position of the hyperplane, and the distance between the hyperplane and the support vectors is called the margin (Fig. 8.21). The

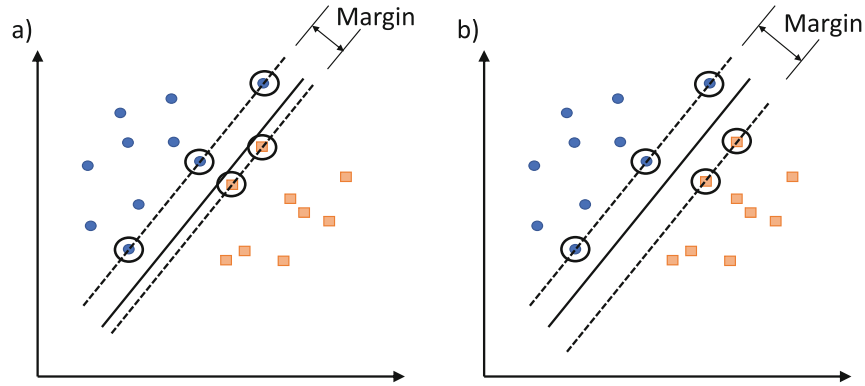



**Fig. 8.21** Examples of support vector machine hyperplanes. (**a**) The margin of the hyperplane is not optimal, (**b**) a hyperplane with a maximized margin. The support vectors are circled

optimal classifier maximizes the margin on both sides of the hyperplane. Because the hyperplane can be defined by only a few of the training samples, SVMs tend to be generalized and robust (Cortes and Vapnik 1995; Duda et al. 2001). When classes cannot be separated linearly, SVMs can map features onto a higher dimensional space where the samples become linearly separable (see Fig. 8.26 in Zeppelzauer et al. 2015).

SVMs originally were designed for binary classification, but a number of methods have been developed for applying them to multi-class problems. The three most common methods are: (1) form $k$ binary "one-against-the-rest" classifiers, where $k$ is the number of classes and the class whose decision-function is maximized is chosen (Vapnik 1998), (2) form all $k(k-1)/2$ pair-wise binary classifiers, and choose the class whose pair-wise decision-functions are maximized (Li et al. 2002), and (3) reformulate the objective function of SVM for the multi-class case so decision boundaries for all classes are optimized jointly (Guemeur et al. 2000).

Gingras and Fitch (2013) used four different algorithms (SVM, k-nearest neighbor, multivariate Gaussian distribution classifier, and GMM) to classify advertisement calls from four genera of anurans and obtained comparable accuracy levels from all three models. Fagerlund (2007) used SVMs to classify bird sounds produced by several species using decision trees with binary SVM classifiers at each node. The two datasets used by Fagerlund (2007) contained six and eight bird species and correct classification scores were 78–88% and 96–98% for the two datasets, respectively, depending on which variables were used in the classifiers.

Zeppelzauer et al. (2015) and Stoeger et al. (2012) both used SVM to identify African elephant rumbles. Zeppelzauer et al. (2015) used cepstral feature vectors and an SVM to distinguish African elephant rumbles from background noise. This SVM resulted in an 88% correct detection rate and a 14% false alarm rate. In addition to SVM, Stoeger et al. (2012) also used linear discriminant analysis (LDA) and nearest neighbor classification algorithms to categorize two types of rumbles produced by five captive African elephants based on spectral representations of the sounds. They obtained a classification accuracy of greater than 97% for all three classification methods.

Jarvis et al. (2006) developed a new type of multi-class SVM, called the class-specific SVM (CS-SVM). In this method, $k$ binary SVMs are created, where each SVM discriminates between one of the $k$ classes of interest and a common reference-class. The class whose decision-function is maximized with respect to the reference-class is selected. If all decision-functions are negative, the reference-class is selected. The advantage of this method is that noise in recordings is treated as the reference-

class. Jarvis et al. (2006) used their CS-SVM to discriminate clicks produced by Blainville's beaked whales from ambient noise and obtained a correct classification score of 98.5%. They also created a multi-class CS-SVM that classified clicks produced by Blainville's beaked whales, spotted dolphins (*Stenella attenuata*), and human-made sonar pings. This CS-SVM resulted in 98% correct classification for Blainville's beaked whale clicks, 88% correct classification for spotted dolphin clicks, and 95% correct classification for sonar pings. It is important to note that the training data were included in their test data, which likely resulted in inflated correct classification scores.

#### 8.4.3.5 Dynamic Time-Warping

Dynamic time-warping (DTW) is a class of algorithms originally developed for automated human speech recognition (Myers et al. 1980). DTW is used to quantitatively compare time-frequency contours of different durations using variable extension and compression of the time axis (Deecke and Janik 2006; Roch et al. 2007). There are different DTW techniques (e.g., Itakura 1975; Sakoe and Chiba 1978; Kruskal and Sankoff 1983), but all are based on comparing a reference sound to a test sound. The test sound is stretched and compressed along its contour to minimize the difference between the shapes of the two contours. Restrictions can be placed on the amount of time-warping that takes place. For example, Buck and Tyack (1993) did not time-warp contours that differed by a factor of more than 2 in duration and assigned those contours a similarity score of zero. Deecke and Janik (2006) stated that contours could only be stretched or compressed up to a factor of 3 to fit the reference contour. In a DTW analysis, all individual contours are compared to all other contours and a similarity matrix is constructed. Sounds are clustered into categories based on the similarity matrix using methods such as k-nearest neighbor cluster analysis or ANNs (Deecke and Janik 2006; Brown and Miller 2007).

DTW has been used to classify bird sounds. Anderson et al. (1996) applied DTW to recognize individual song syllables for two species of songbirds: indigo buntings (*Passerina cyanea)* and zebra finches (*Taeniopygia guttata*). Their analysis resulted in 97% correct classification of stereotyped syllables and 84% correct classification of syllables in plastic song. It is important to note, however, that these results were obtained for song recorded from a single individual of each species in a controlled setting. Somervuo et al. (2006) performed DTW to classify bird song syllables produced by 14 different species. They compared two different methods for computing distance between syllables: (1) simple Euclidean distances between frequency-amplitude vectors, and (2) absolute distance between frequencies weighted by the sum of their amplitudes. Classification accuracy was low, at about 40–50%, depending on the species and the distance method used. They obtained higher classification success using classification methods such as hidden Markov models (HMM) and GMM based on song fragments, rather than on single syllables.

Buck and Tyack (1993) performed DTW to classify three signature whistles from each of five wild bottlenose dolphins recorded in Sarasota, Florida, USA, with 100% accuracy. Deecke and Janik (2006) used DTW to classify signature whistles produced by captive bottlenose dolphins. The DTW algorithm outperformed human analysts and other statistical methods tested by Janik (1999). DTW also was applied to classify stereotypical pulsed sounds produced by killer whales, both in captivity (Brown et al. 2006) and at sea (Deecke and Janik 2006; Brown and Miller 2007). In all of these studies, sounds were classified into categories that were identified perceptually by humans with very high correct classification scores.

Oswald et al. (2021) used dynamic time-warping and neural network analysis to group whistle contours produced by short- and long-beaked common dolphins (*Delphinus delphis* and *D. bairdii*) into categories. Many of the resulting categories were shared between the two species, but each species also produced a number of species-specific categories. Random forest analysis showed that whistles in species-specific categories could be classified to species with significantly higher accuracy than whistles

in shared categories. This suggests that not every whistle carries species information, and that specific whistle types play an important role in dolphin species identification.

#### 8.4.3.6 Hidden Markov Models

Hidden Markov mode (HMM) theory was developed in the late 1960s by Baum and Eagon (1967) and now is used commonly for human speech recognition (Rabiner et al. 1983, 1996; Levinson 1985; Rabiner 1989). To create an HMM, a vector of features is extracted from a signal at discrete time steps. The temporal evolution of these features from one state to the next is modeled by creating a transition matrix $M$, where $M_{ij}$ is the probability of transition from state $i$ to state $j$, and an emission matrix $E$, where $E_{\text{is}}$ is the probability of observing signal $s$ in state $i$ (Rickwood and Taylor 2008). A different HMM is created for each species in the dataset and a sound is classified by determining which of the HMMs has the highest likelihood of producing that particular set of signal states. Training HMMs requires significant amounts of computing, and proper estimation of the transition and output probabilities is of crucial importance (Makhoul and Schwarz 1995). Excellent tutorials on HMMs can be found in Rabiner and Juang (1986) and Rabiner (1989).

A significant advantage inherent to HMMs is their ability to model time and spectral variability simultaneously (Makhoul and Schwarz 1995). They are able to model time series that have subtle temporal structure and are efficient for modeling signals with varying durations by performing nonlinear, temporal alignment during both the training and classification processes (Clemins et al. 2005; Roch et al. 2007; Trifa et al. 2008). Using HMMs, complex models can be built to deal with complicated biological signals (Rickwood and Taylor 2008), but care must be taken when choosing training samples to obtain a high generalization ability. The performance of an HMM is influenced by the size of the training set, the feature extraction method, and the number of states in the model (Trifa et al. 2008). Recognition performance is also affected by noise (Trifa et al. 2008).

In addition to being successfully implemented in human speech recognition, HMMs have been used to classify the sounds produced by birds (Kogan and Margoliash 1998; Trawicki et al. 2005, Trifa et al. 2008, Adi et al. 2010), red deer (*Cervus elaphus*; Reby et al. 2006), African elephants (Clemins et al. 2005), common dolphins (Sturtivant and Datta 1997; Datta and Sturtivant 2002), killer whales (Brown and Smaragdis 2008, 2009); beluga whales (Clemins and Johnson 2005; Leblanc et al. 2008), bowhead whales (Mellinger and Clark 2000), and humpback whales (Suzuki et al. 2006). HMMs perform as well as, or better than, both GMMs and DTW (Weisburn et al. 1993; Kogan and Margoliash 1998) and are becoming more common in animal classification studies.

Adi et al. (2010) also used HMMs to examine individually distinct acoustic features in songs produced by ortolan buntings (*Emberiza hortulana*). They represented each song syllable using a 15-state HMM (Fig. 8.22). These HMMs then were connected to represent song types. The 14 most common song types were included in the analysis and correct classification ranged from 50% to 99%, depending on the song type. Overall, 90% of songs were correctly classified. Adi et al. (2010) used these results to illustrate the feasibility of using acoustic data to assess population sizes for these birds.

Reby et al. (2006) used HMMs to examine whether common roars uttered by red deer during the rutting season can be used for individual recognition. They recorded roar bouts from seven captive red deer and used HMMs to model roar bouts as successions of silences and roars. Each roar in the analysis was modeled as a succession of states of frequency components measured from the roars. Overall, the HMM correctly identified 85% of roar bouts to the individual deer, showing that roars were individually specific. Reby et al. (2006) also used HMMs to examine stability in this individuality over the rutting season. They did this by training an HMM using roar bouts recorded at the beginning of the rutting season and testing the model using roar bouts recorded later in the rutting season. Overall, 58% of roar bouts were classified correctly, suggesting that individual identification cues in roar bouts varied over time.

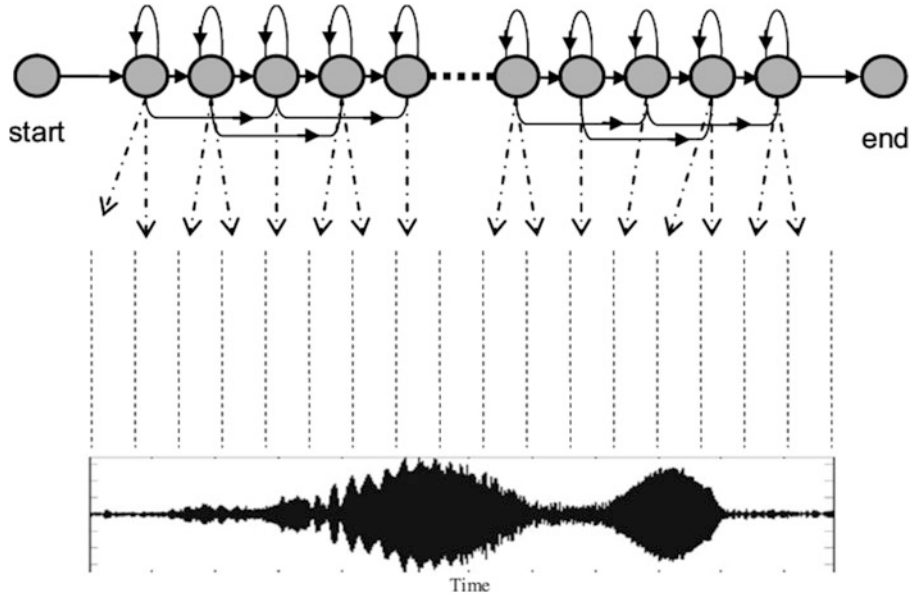



**Fig. 8.22** Example of a 15-state hidden Markov model representation of the waveform of a song syllable produced by an ortolan bunting to capture the temporal pattern of the syllable (Adi et al. 2010). © Acoustical Society of America, 2010. All rights reserved

## 8.5 Challenges in Classifying Animal Sounds

Placing sounds into categories is not always straightforward. Sounds produced by a particular species often contain a great deal of variability caused by different factors (e.g., location, date, age, sex, and individuality), which can make it difficult to define categories. In addition, sound categories are not always sharply demarcated, but instead grade or gradually transition from one form to another. It is important to be aware of the challenges in a particular dataset. Below are some types of variation that can be encountered in the classification of animal sounds.

### 8.5.1 Recording Artifacts

Bioacousticians need to be aware that recorded animal sounds are affected by the frequency and sensitivity specifications of the recording system used. An inappropriate recording system can result in distorted or partial sounds, which complicates their classification. For example, sounds can be misrepresented in recordings if the frequency response of the recording system is not linear, if the sampling frequency is too low, if sounds exist below or above the functional frequency range of the recording system, or if aliasing occurs (see Chap. 4). Ideally, recording systems should be carefully assembled and calibrated for the specific application. If the effects of the recording system could always be removed completely from recordings, sound classification would be more consistent and comparable. However, sounds published in the literature are sometimes received sounds that were affected by the recorder and/or the sound propagation environment.

One of the most common problems in underwater acoustic recordings is mooring noise. If hydrophones are held over the side of a boat, the recordings will contain sound from waves splashing against the boat or the hydrophone cable rubbing against the boat. Recorders built into mooring lines can record cable strum or clanking chains. If multiple oceanographic sensors are moored together, sounds from other

instruments (e.g., wipers on a turbidity sensor) may be recorded. Recorders resting on soft seafloor in coastal water may record the sound of sand swishing over the mooring. In addition, hydrostatic pressure fluctuations from the recorder bouncing in the water column or vortices at the hydrophone if deployed in strong currents will cause flow noise. All of these artifacts can last from seconds to minutes and appear in spectrograms as power from a few hertz to high kilohertz. Minimization of mooring noise and identification of recording artifacts is an art (also see Chaps. 2 and 3).

Similarly, artifacts can be recorded during airborne recordings. Wind is a primary artifact; however, moving vegetation and precipitation can also add noise to a recording. Any disturbance to the microphone can generate unwanted tapping or static on a recording. Recording systems in terrestrial environments need to be secured to minimize such noises.

### 8.5.2 Sound Propagation Effects

Environmental features of air or water can change the way sound propagates and thus the acoustic characteristics of a recorded sound. Bioacousticians need to understand environmental effects on the features of received sound to avoid classification of a signal variant as a new type, rather than as a particular sound type affected by propagation conditions. The sound propagation environment can affect both the spectral and temporal features of sound as it propagates from the animal to the recorder (see Chaps. 5 and 6). For example, energy at high frequencies is lost (attenuates) very quickly due to scattering and absorption, and therefore high-frequency harmonics do not propagate over long ranges. Acoustic energy at low frequencies (i.e., long wavelengths) does not travel well in narrow waveguides (e.g., shallow water). Because different frequencies within a sound can attenuate at different rates, the same sound can appear differently on a spectrogram, depending on the distance at which it was recorded.

Differential attenuation of frequencies in air is shown in Fig. 8.23. Signals produced by a big brown bat (*Eptesicus fuscus*) flying toward a microphone contain more ultrasonic components than signals recorded from a bat flying away from the microphone. The signal with the longest frequency modulation (from 100 to 50 kHz) is received when the bat is closest to the microphone. Variations in this spectrogram show how one sound type could be categorized differently simply because of distance between the animal and recorder, orientation to the microphone, and the gain setting.

Other sound propagation effects include reverberation (which leads to the temporal spreading of brief, pulsed sounds) and frequency dispersion. Frequency dispersion is a result of energy at different frequencies traveling at different speeds. This leads to sounds being spread out in time and, specifically in some underwater environments, can cause pulsed sounds to become frequency-modulated sounds (either up- or downsweeps; Fig. 8.24).

Finally, ambient noise (i.e., geophysical noise, anthropogenic noise, and non-target biological noise) superimposes with animal sounds, and at some distances and frequencies, parts of the animal sound spectrum will begin to drop below the levels of ambient noise. As a result, the same animal sound in a different environment and at a different distance from the animal can look quite different on a spectrogram and cause it to be misclassified as two different sound types.

### 8.5.3 Angular Aspects of Sound Emission

The orientation of an animal relative to the receiver (microphone or hydrophone) can change the acoustic features of the recorded sound. This complicates classification, and off-axis variations of a sound need to be known so they can be categorized as just a variant of a particular sound type, rather than as a new sound type. Not all sounds emitted by animals are omnidirectional (i.e., propagate equally in all angles relative to the animal). Au et al. (2012) studied the directionality of bottlenose dolphin echolocation clicks by measuring the horizontal and vertical emission beam patterns of these sounds. The angle at which an echolocation click was

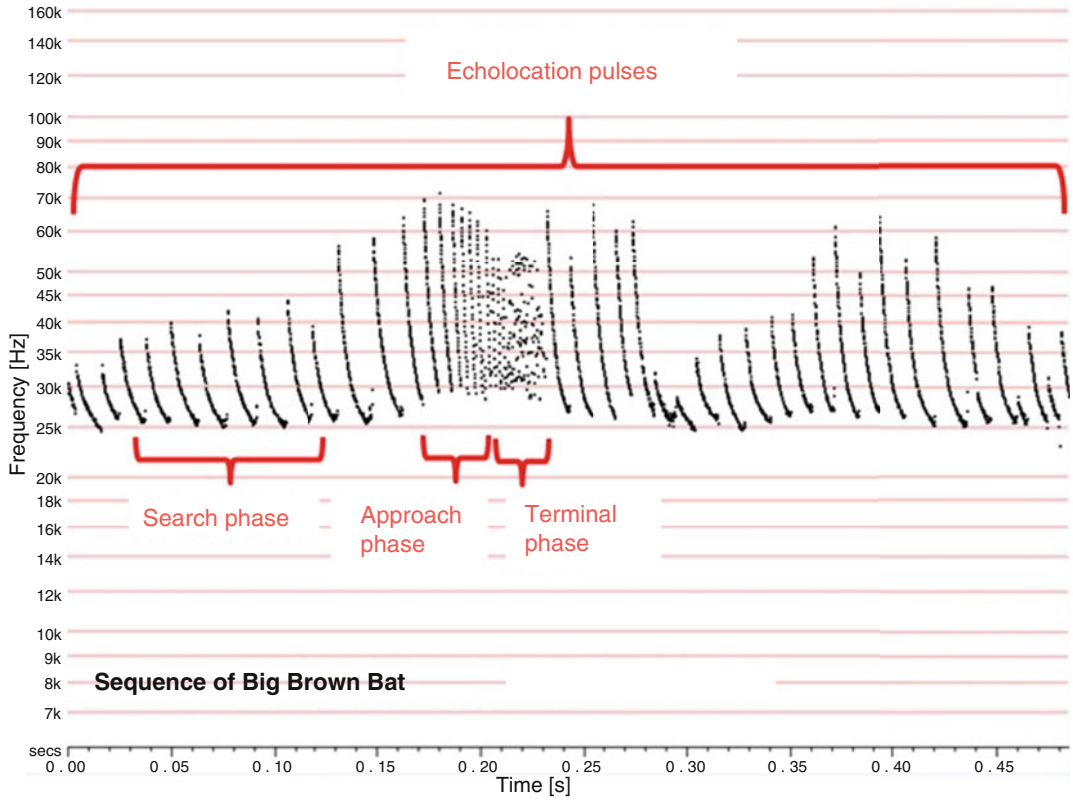



**Fig. 8.23** Spectrogram of big brown bat (*Eptesicus fuscus*) circling a recording device while searching and pursuing aerial prey. As the bat approaches the microphone, more of the ultrasonic signal is received (calls reach up to 70 kHz). As the bat moves away, the signal is attenuated. Time between calls shortens notably as the bat pursues an insect prey for capture. Notice that the bat emits "search" calls at 25–40 kHz, approach calls at 30–70 kHz when it is in pursuit or trying to navigate flight through complex space, and finally terminal calls at 30–55 kHz

recorded relative to the transducer (or echolocating animal) not only affected its received level, but also the waveform and frequency spectrum (Fig. 8.25). Sperm whale (*Physeter macrocephalus*) echolocation clicks, when recorded off-axis (i.e., away from the center of its emission beam), consisted of multiple complex pulses that were likely due to internal reflections within the sperm whale's head (Møhl et al. 2003; also see Chap. 12).

### 8.5.4 Geographic Variation

Geographic variation, or differences in the sounds produced by populations of the same species living in different regions, has been documented for many terrestrial and aquatic animals, including Hawaiian crickets (Mendelson and Shaw 2003), Túngara frogs (*Engystomops pustulosus*, Pröhl et al. 2006), bats (Law et al. 2002; Aspetsberger et al. 2003; Russo et al. 2007; Yoshino et al. 2008), pikas (Borisova et al. 2008), sciurid rodents (Gannon and Lawlor 1989; Slobodchikoff et al. 1998; Yamamoto et al. 2001; Eiler and Banack 2004), singing mice (*Scotinomys* spp., Campbell et al. 2010), primates (Mitani et al. 1992; Delgado 2007; Wich et al. 2008), cetaceans (Helweg et al. 1998; McDonald et al. 2006; Delarue et al. 2009; Papale et al. 2013, 2014), and elephant seals (*Mirounga* spp., Le Boeuf and Peterson

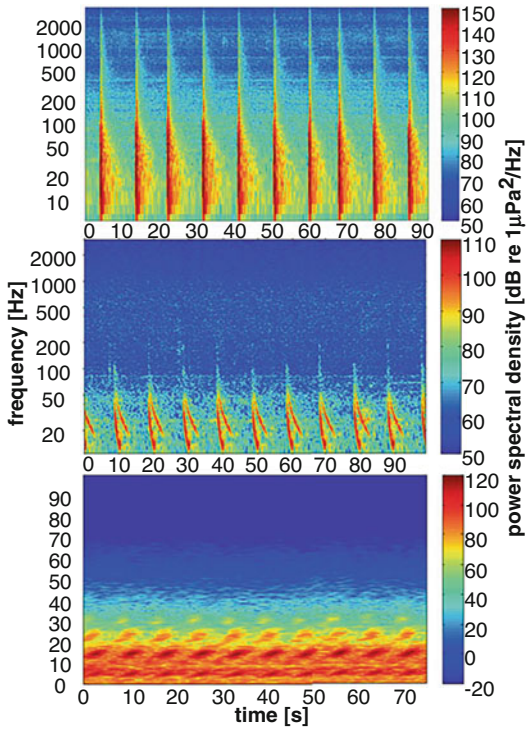



**Fig. 8.24** Spectrograms of marine seismic airgun signals recorded at three different ranges: 1.5 km (top), 80 km over soft seabed (middle), and 40 km over a hard seabed (bottom). The top and bottom spectrograms are of the same seismic survey. Pulses were brief and broadband near the source, but became frequency-modulated and narrowband some distance away due to dispersion (Erbe et al. 2016). © Erbe et al.; https://ars.els-cdn.com/content/image/1-s2.0-S0025326X15302125-gr9_lrg.jpg. Licensed under CC BY 4.0; https://creativecommons.org/licenses/by/4.0/

1969). When developing classifiers, it is important to understand the degree of geographic variation in a sound repertoire and the range over which this occurs. If geographic variation exists, then a classifier trained using data collected in one location may not work well when applied to data collected in another location.

One of the underlying causes of geographic variation may be reproductive isolation of a population. Keighley et al. (2017) used DFA with stepwise variable selection to determine geographic variation in sounds from six major populations of palm cockatoos (*Probosciger aterrimus*) in Australia. Palm cockatoos from the east coast (Iron Range National Park) had unique contact sounds and produced fewer sound types than at other locations. The authors speculated that this large difference was due to long-term isolation at this site and noted that documentation of geographic variation in sounds provided important conservation information for determining connectivity of these six populations.

Thomas and Golladay (1995) employed PCA to classify nine underwater vocalization types produced by leopard seals (*Hydrurga leptonyx*) at three study sites near Palmer Peninsula, Antarctica. The PCA successfully separated vocalizations from the three study areas and provided information about what features of the sounds were driving the differences among locations. For example, the first principal component was influenced by maximum, minimum, start, and end frequencies, the second principal component was influenced by the presence or absence of overtones, and the third principal component was predominantly related to time relationships, such as duration and time between successive sounds. Note that some sound types were absent at some locations.

### 8.5.5 Graded Sounds

Some animals produce sound types that grade or gradually transition from one type to another. Researchers should not neglect the potential existence of vocal intermediates in classification. For example, Schassburger (1993) described sounds produced by timber wolves (*Canis lupus*) as barks, growl-moans, growls, howls moans, snarls, whimpers, whine-moans, whines, woofs, and yelps. Wolves combine these 11 principal sounds to create mixed-sounds that often grade from one type into another.

Clicks trains, burst-pulse sounds, and whistles produced by delphinids are typically considered as three distinct categories of sound. Click trains and burst-pulse sounds are composed of short, exponentially damped sine waves separated by periods of silence, while whistles are generally thought of as continuous tonal sounds, often

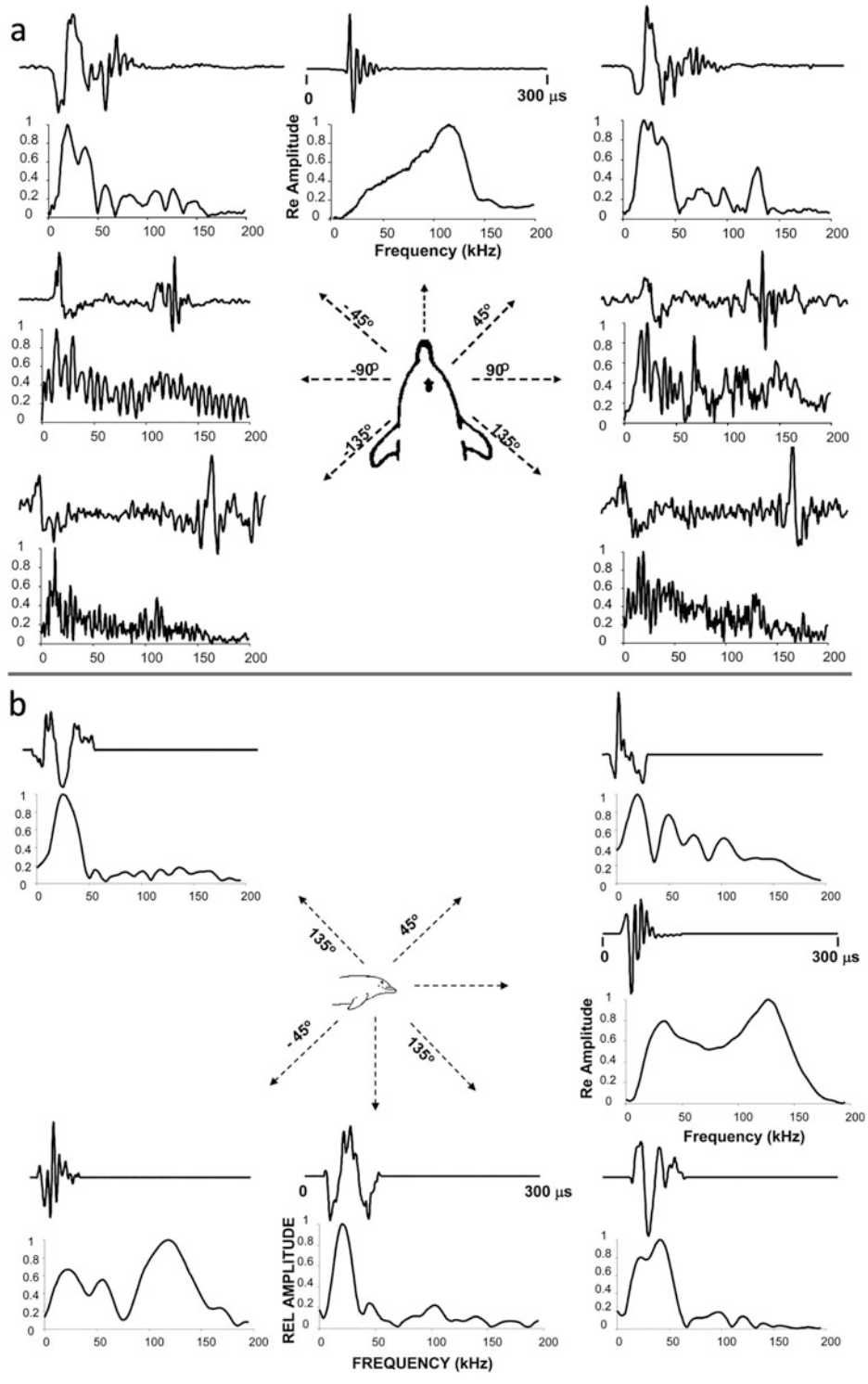



**Fig. 8.25** Waveforms and spectra of a bottlenose dolphin echolocation click in the horizontal (**a**) and vertical (**b**) planes (Au et al. 2012). © Acoustical Society of America, 2012. All rights reserved

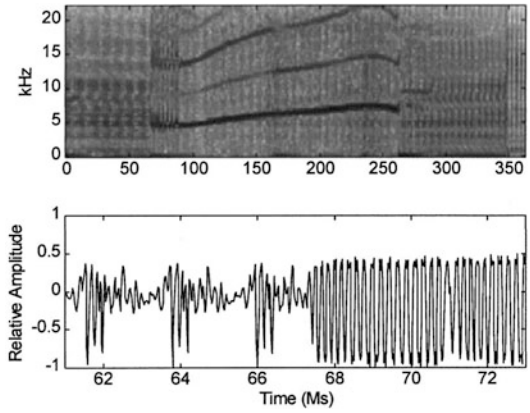



**Fig. 8.26** Spectrogram and waveform of a false killer whale vocalization. The vocalization appears to be a whistle in the spectrogram, but the waveform reveals discrete pulses between 61 and 67 ms (Murray et al. 1998). © Acoustical Society of America, 1998. All rights reserved

sweeping in frequency. While these sounds appear quite different from one another on spectrograms, closer inspection of their waveforms reveals that some sounds that look like whistles on a spectrogram actually contain a high degree of amplitude modulation. In other words, some sounds that are considered to be whistles are made up of pulses with inter-pulse intervals that are too short to hear or be resolved by the analysis window of the spectrogram (Fig. 8.26). As an example of this, Murray et al. (1998) used self-organizing neural networks to analyze the vocal repertoires of two captive false killer whales (*Pseudorca crassidens*) based on measurements taken from waveforms. They found that rather than organizing sounds into distinct categories, the vocal repertoire was more accurately represented by a graded continuum, with exponentially damped sinusoidal pulses on one end and continuous sinusoidal signals at the other. Beluga whales also have been shown to have a graded vocal repertoire (Karlsen et al. 2002; Garland et al. 2015). Whistles with a high degree of amplitude modulation have been recorded from Atlantic spotted and spinner (*Stenella longirostris*) dolphins (Lammers et al. 2003), suggesting that this graded continuum model is applicable to these species as well.

### 8.5.6 Repertoire Changes Over Time

Some animal sound repertoires change over time, which complicates their classification. For example, humpback whale song slowly changes over the course of a breeding season as new units are introduced and old ones discarded (Noad et al. 2000). Song also changes from one season to the next, and in one instance, eastern Australian humpback whales changed to the song of the western Australian population within 1 year (Noad et al. 2000).

Antarctic blue whales can be heard off southwestern Australia from February to October every year. The upper frequency of their Z-call decreases over the season by about 0.4–0.5 Hz. At the beginning of the next season, the Z-call jumps in frequency to about the mean of the Z frequency of the previous season, and then decreases again, leading to an average decrease in the frequency of the upper part of the Z-call by $0.135 \pm 0.003$ Hz/year (Fig. 8.27; Gavrilov et al. 2012). A similar decrease (albeit at different rates at different locations) has been observed for the "spot call," of which the animal source remains elusive (Fig. 8.27; Ward et al. 2017). The reasons for these shifts are unknown.

## 8.6 Summary

Animals, whether they are in air, on land, or under water, produce sound in support of their various life functions. Cicadas join in chorus to repel predatory birds (Simmons et al. 1971); male fishes chorus on spawning grounds to attract females (Amorim et al. 2015); frogs call to attract mates and to mark out their territory (Narins et al. 2006); birds, too, sing for territorial and reproductive reasons (Catchpole and Slater 2008); bats emit clicks for echolocation during hunting and navigating, as do dolphins (Madsen and Surlykke 2013). In order to study animals by listening to their sounds, sounds need to be classified to species, to behavior, etc. In the early days, this was done without measurements or with only the simplest measuring tools. Scientists listened to the

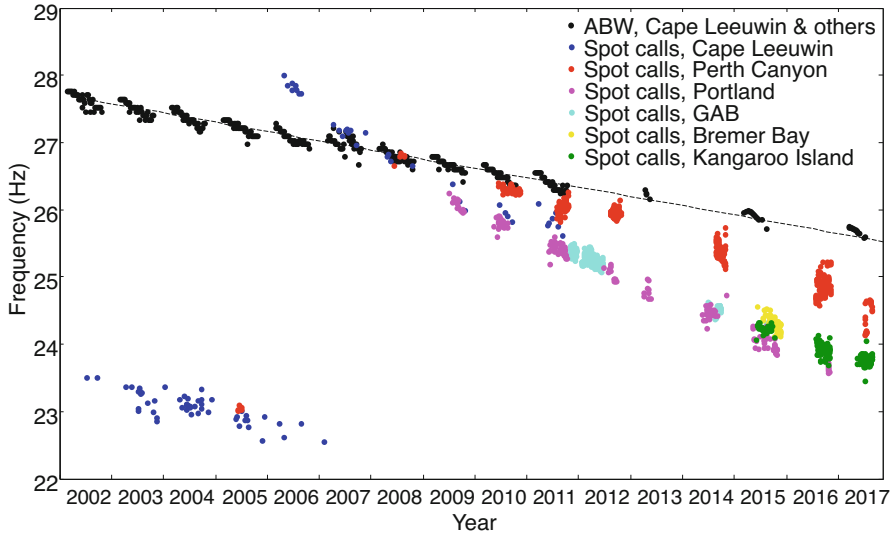



**Fig. 8.27** Weekly means of the upper part of the Antarctic blue whale Z-call over several years, as well as of the spot call, which remains to be identified to species. All locations are off Australia (GAB: Great Australian Bight). Data updated from Gavrilov et al. (2012) and Ward et al. (2017). Courtesy of Sasha Gavrilov

sounds in the field, often while visually observing animals. Scientists recorded sounds in the field and analyzed the recordings in the laboratory by listening, looking at oscillograms or spectrograms, and manually sorting sounds into types. Nowadays, with the affordability of autonomous recording equipment, bioacousticians collect vast amounts of data, which can no longer be analyzed without the aid of automated data processing, data reduction, and data analysis tools. Given simultaneous advances in computer hard- and software, datasets may be analyzed more efficiently, and with the added advantage of reducing opportunities for human subjective biases.

In this chapter, we presented software tools for automatically detecting animal sounds in acoustic recordings, and for classifying those sounds. The detectors we discussed compute a specific quantity of the sound (such as its instantaneous energy or entropy) and then apply a threshold above which the sound is deemed detected. The specific detectors were based on acoustic energy, Teager–Kaiser energy, entropy, matched filtering, and spectrogram cross-correlation. Setting the detection threshold critically affects how many signals are detected and how many are missed. We presented two ways of finding the best threshold and assessing detector performance: receiver operating characteristics and precision-recall curves.

Once signals have been detected, they can be classified. A common pre-processing step immediately prior to classification includes the measurement of sound features such as minimum and maximum frequency, duration, or cepstral features. The software tools we presented for classification included parametric clustering, principal component analysis, discriminant function analysis, classification trees, and machine learning algorithms. No single tool outperforms all others; rather, the best tool suited for the specific task needs to be employed. We discussed advantages and limitations of the various tools and provided numerous examples from the literature. Finally, challenges resulting from recording artifacts, the environment affecting sound features, and changes in sound features over time and space were explored.

It is important to remember that human perception of a sound likely is not the same as an animal's perception of the sound and yet

bioacousticians commonly describe or classify animal sounds in human terms. Classification of the acoustic repertoire of an animal into sound types provides a convenient framework for comparing and contrasting sounds, taking systematic measurements from portions of the repertoire, and performing statistical analyses. However, categories determined based on human perception may have little or no relevance to the animals and so human categorizations can be biologically meaningless. For example, humans have limited low-frequency and high-frequency hearing abilities compared to many other species, and so aural classification of sound types is sometimes based on only a portion of a sound audible to the human listener. Whether sound types determined by humans are meaningful classes to the animals is mostly unknown. While categorizing sounds based on function is an attractive approach for the behavioral zoologist, establishing the functions of these sounds is often challenging. In our review of classification methods, it was clear that methods developed for human speech could be applied to animal sounds. Some fascinating questions lie ahead for bioacousticians as they attempt to extend understanding of the perception experienced by other animals.

Even with the above caveats, detection and classification of animal sounds is useful for research and conservation. It allows populations to be monitored, their distribution and abundance to be determined, and impacts (e.g., from human presence or climate change) to be assessed. It can also be useful for conservation of a species (i.e., to create taxonomy, identify geographic variation in populations, examine ecological connectivity among populations, and detect changes in the biological uses sounds due to the advent and growth of anthropogenic noise). Classification of animal sounds is important for understanding behavioral ecology and social systems of animals and can be used to identify individuals, social groups, and populations. The ability to study these types of topics will ultimately lead to a deeper understanding of the evolutionary forces that shape animal bioacoustics.

With a goal to foster wider participation in research on bioacoustic pattern recognition, a number of global competitions are held regularly. The annual Detection and Classification of Acoustic Scenes and Event (DCASE) workshops and BirdCLEF challenges (part of Cross Language Evaluation Forum) attract hundreds of data scientists for developing machine learning solutions for recognizing bird sounds in soundscape recordings. The marine mammal community organizes the biennial Detection, Classification, Localization, and Density Estimation (DCLDE) workshops. These challenges put out large training datasets for researchers to develop detection and classification systems, assess the performance of submitted solutions with "held out" datasets, and reward the top-ranked submissions. The datasets from these challenges are often made available for use by the research community after the competitions, while some workshops make available the submitted solutions as well.

## 8.7 Additional Resources

- *PAMGuard* is an open-source software package for acoustic detection, classification, and localization of cetacean sounds: https://www.pamguard.org/
- *Ishmael* is a free software package for acoustic detection, classification, and localization of cetacean sounds: http://www.bioacoustics.us/ishmael.html
- *Koe* is a free, web-based software for annotation, measurement, and classification of bioacoustics signals: https://koe.io.ac.nz/# (Fukuzawa et al. 2020)
- *Praat* is free software originally designed for human speech analysis, but used by many bioacousticians: https://www.fon.hum.uva.nl/praat/
- *Characterization Of Recorded Underwater Sound (CHORUS)* is a *MATLAB* graphic user interface developed by Curtin University, Perth, WA, Australia, with built-in automatic detectors for pygmy blue and fin whales

(Gavrilov and Parsons 2014): https://cmst.curtin.edu.au/products/chorus-software/
- Detection, Classification, Localization, and Density Estimation of Marine Mammals using Passive Acoustics meeting websites:
  - Mount Hood, Oregon, USA, 2011: http://www.bioacoustics.us/dcl.html
  - St Andrews, Scotland, UK, 2013: https://soi.st-andrews.ac.uk/dclde2013/
  - San Diego, California, USA, 2015: http://www.cetus.ucsd.edu/dclde/index.html
  - Paris, France, 2018: http://sabiod.univ-tln.fr/DCLDE/
  - Hawaii, USA, 2022: http://www.soest.hawaii.edu/ore/dclde/
- Bird sound recognition challenges: http://dcase.community/ (DCASE), https://www.imageclef.org/BirdCLEF2020 (BirdCLEF)
- *BirdNET* is an Android app for birdsong recognition: https://birdnet.cornell.edu/
- *SongSleuth* is an Apple or Android app for birdsong recognition: https://www.songsleuth.com/#/
- All accessed 5 Aug 2022.

## References


Adi K, Johnson MT, Osiejuk TS (2010) Acoustic censusing using automatic vocalization classification and identity recognition. J Acoust Soc Am 127:874–883. https://doi.org/10.1121/1.3273887

Afifi AA, Clark V (1996) Computer-aided multivariate analysis, 3rd edn. Chapman and Hall/CRC, New York

Amorim MC, Vasconcelos RO, Fonseca PJ (2015) Fish sounds and mate choice. In: Ladich F (ed) Sound communication in fishes. Springer, Vienna, pp 1–33

Anderson SE, Dave AS, Margoliash D (1996) Template-based automatic recognition of birdsong syllables from continuous recordings. J Acoust Soc Am 100:1209–1219. https://doi.org/10.1121/1.415968

Armitage DW, Ober HK (2010) A comparison of supervised learning techniques in the classification of bat echolocation calls. Ecol Inform 5:465–473. https://doi.org/10.1016/j.ecoinf.2010.08.001

Aspetsberger F, Brandsen D, Jacobs DS (2003) Geographic variation in the morphology, echolocation and diet of the little free-tailed bat, *Chaerephon pumilus* (Molossidae). Afr Zool 38:245–254. https://doi.org/10.1080/15627020.2003.11407278

Au WWL, Nachtigall PE (1995) Artificial neural network modeling of dolphin echolocation. In: Kastelein RA, Thomas JA, Nachtigall PE (eds) Sensory systems of aquatic mammals. De Spil Publishers, Woerden, The Netherlands, pp 183–199

Au WWL, Branstetter B, Moore P, Finneran J (2012) The biosonar field around an Atlantic bottlenose dolphin (*Tursiops truncatus*). J Acoust Soc Am 131(1): 569–576. https://doi.org/10.1121/1.3662077

Baptista LF, Gaunt SSL (1997) Social interaction and vocal development in birds. In: Snowden CT, Hausberger M (eds) Social influences on vocal development. Cambridge Univ Press, Cambridge, pp 23–40

Baum LE, Eagon JA (1967) An inequality with applications to statistical estimation for probabilistic functions of Markov processes and to a model for ecology. Bull Am Math Soc 73:360–363

Baumgartner MF, Fratantoni DM (2008) Diel periodicity in both Sei whale vocalization rates and the vertical migration of their copepod prey observed from ocean gliders. Limnol Oceanogr 53:2197–2209. https://doi.org/10.4319/lo.2008.53.5_part_2.2197

Beeman K (1998) Digital signal analysis, editing and synthesis. In: Hopp SL, Owren MJ, Evans CS (eds) Animal acoustic communication: sound analysis and research methods. Springer, Berlin, pp 59–103

Belliustin NS, Kuznetsov SO, Nuidel IV, Yakhno VG (1991) Neural networks with close nonlocal coupling for analyzing composite image. Neurocomputing 3: 231–246. https://doi.org/10.1016/0925-2312(91)90005-V

Bergler C, Schröter H, Cheng RX, Barth V, Weber M, Nöth E, Hofer H, Maier A (2019) ORCA-SPOT: an automatic killer whale sound detection toolkit using deep learning. Sci Rep 9(1):1–7. https://doi.org/10.1038/s41598-019-47335-w

Bermant PC, Bronstein MM, Wood RJ, Gero S, Gruber DF (2019) Deep machine learning techniques for the detection and classification of sperm whale bioacoustics. Sci Rep 9(1):1–10. https://doi.org/10.1038/s41598-019-48909-4

Borisova NG, Rudneva LV, Starkov AI (2008) Interpopulation variability of vocalizations in the Daurian pika (*Ochotona daurica*). Zool Zh 87:850–861

Bouffaut L, Dréo R, Labat V, Boudraa AO, Barruol G (2018) Passive stochastic matched filter for Antarctic blue whale call detection. J Acoust Soc Am 144(2): 955–965. https://doi.org/10.1121/1.5050520

Bradbury JW, Vehrencamp SL (2011) Principles of animal communication, 2nd edn. Sinauer Associates, New York

Brandes TS (2008) Feature-vector selection and use with Hidden Markov Models to identify frequency-modulated bioacoustic signals amidst noise. IEEE Trans Speech Lang Process 16:1173–1180. https://doi.org/10.1109/TASL.2008.925872

Breiman L (2001) Random forests. Mach Learn 45:5–32

Breiman L, Friedman J, Olshen R, Stone C (1984) Classification and regression trees. Wadsworth, Pacific Grove, CA

Briefer EF, Maigrot A-L, Roi T, Mandel R, Briefer Freymond S, Bachmann I, Hillmann E (2015) Segregation of information about emotional arousal and

valence in horse whinnies. Sci Rep 5(1):1–11. https://doi.org/10.1038/srep09989
Briskie JV, Martin PR, Martin TE (1999) Nest predation and the evolution of nestling begging calls. Proc R Soc Lond B 266:2153–2159. https://doi.org/10.1098/rspb.1999.0902
Brown JC, Miller PJO (2007) Automatic classification of killer whale vocalizations using dynamic time warping. J Acoust Soc Am 122:1201–1207. https://doi.org/10.1121/1.2747198
Brown JC, Smaragdis P (2008) Automatic classification of vocalizations with Gaussian mixture models and Hidden Markov Models. J Acoust Soc Am 123:3345. https://doi.org/10.1121/1.2933896
Brown JC, Smaragdis P (2009) Hidden Markov and Gaussian mixture models for automatic sound classification. J Acoust Soc Am 125:EL221–EL224. https://doi.org/10.1121/1.3124659
Brown JC, Hodgins-Davis A, Miller PJO (2006) Classification of vocalizations of killer whales using dynamic time warping. J Acoust Soc Am 119:EL34–EL40. https://doi.org/10.1121/1.2166949
Buck JR, Tyack PL (1993) A quantitative measure of similarity for *Tursiops truncatus* signature whistles. J Acoust Soc Am 94:2497–2506. https://doi.org/10.1121/1.407385
Camacho-Alpízar A, Fuchs EJ, Barrantes G (2018) Effect of barriers and distance on song, genetic, and morphological divergence in the highland endemic Timberline Wren (*Thryorchilus browni*, Troglodytidae). PLoS One 13(12):e0209508. https://doi.org/10.1371/journal.pone.0209508
Campbell P, Pasch B, Pino JL, Crino OL, Phillips M, Phelps SM (2010) Geographic variation in the songs of neotropical singing mice: testing the relative importance of drift and local adaptation. Evolution 64(7):1955–1972. https://doi.org/10.1111/j.1558-5646.2010.00962.x
Catchpole CK, Slater PJB (2008) Bird song: biological themes and variations, 2nd edn. Cambridge University Press, Cambridge
Cerchio S, Jacobsen JK, Norris TF (2001) Temporal and geographical variation in songs of humpback whales, *Megaptera novaeangliae*: synchronous change in Hawaiian and Mexican breeding assemblages. Anim Behav 62(2):313–329. https://doi.org/10.1006/anbe.2001.1747
Cho K, Van Merriënboer B, Bahdanau D, Bengio Y (2014) On the properties of neural machine translation: encoder-decoder approaches. arXiv:1409.1259
Clark CW (1980) A real-time direction-finding device for determining the bearing to the underwater sounds of southern right whales, *Eubalaena australis*. J Acoust Soc Am 68:508–511. https://doi.org/10.1121/1.384762
Clark CW (1982) The acoustic repertoire of the southern right whale, a quantitative analysis. Anim Behav 30(4):1060–1071. https://doi.org/10.1016/S0003-3472(82)80196-6
Clark LA, Pregibon D (1992) Statistical models. In: Chambers SJM, Hastie TJ (eds) Statistical models in S. Wadsworth and Brooks/Cole, Pacific Grove, CA
Clarke E, Reichard UH, Zuberbühler K (2006) The syntax and meaning of wild gibbon songs. PLoS One 1(1):E73. https://doi.org/10.1371/journal.pone.0000073
Clemins PJ, Johnson MT (2005) Unsupervised classification of beluga whale vocalizations. J Acoust Soc Am 117:2470. https://doi.org/10.1121/1.4809461
Clemins PJ, Johnson MT, Leong KM, Savage A (2005) Automatic classification and speaker identification of African elephant (*Loxodonta africana*) vocalizations. J Acoust Soc Am 117:956–963. https://doi.org/10.1121/1.1847850
Coifman RR, Lafon S, Lee AB, Maggioni M, Nadler B, Warner F, Zucker SW (2005) Geometric diffusions as a tool for harmonic analysis and structure definition of data: diffusion maps. Proc Natl Acad Sci 102(21):7426–7431. https://doi.org/10.1073/pnas.0500334102
Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20:273–297
Courts R, Erbe C, Wellard R, Boisseau O, Jenner KC, Jenner M-N (2020) Australian long-finned pilot whales (*Globicephala melas*) emit stereotypical, variable, biphonic, multi-component, and sequenced vocalisations, similar to those recorded in the northern hemisphere. Sci Rep 10(1):20609. https://doi.org/10.1038/s41598-020-74111-y
Crance JL, Berchok CL, Wright DL, Brewer AM, Woodrich DF (2019) Song production by the North Pacific right whale, *Eubalaena japonica*. J Acoust Soc Am 145(6):3467–3479. https://doi.org/10.1121/1.5111338
Cutler DR, Edwards TC Jr, Beard KH, Cutler A, Hess KT, Gibson J, Lawler JJ (2007) Random forests for classification in ecology. Ecology 88:2783–2792. https://doi.org/10.1890/07-0539.1
Dang T, Bulusu N, Hu W (2008) Lightweight acoustic classification for cane toad monitoring. In: 42nd Asilomar Conference on Signals, Systems and Computers. IEEE, New York, pp 1601–1605
Datta S, Sturtivant C (2002) Dolphin whistle classification for determining group identities. Sig Process 82(2):251–258. https://doi.org/10.1016/S0165-1684(01)00184-0
Davis J, Goadrich M (2006) The relationship between precision-recall and ROC curves. In: Proceedings of the 23rd International Conference on Machine Learning, Pittsburgh, PA
Davis SB, Mermelstein P (1980) Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. IEEE Trans Acoust Speech Sig Process 28:357–366. https://doi.org/10.1109/TASSP.1980.1163420
Dawson MRW, Charrier I, Sturdy CB (2006) Using an Artificial Neural Network to classify black-capped chickadee (*Poecile atricapillus*) sound note types. J Acoust Soc Am 119(5):3161–3172. https://doi.org/10.1121/1.2189028

Deecke VB, Janik VM (2006) Automated categorization of bioacoustic signals: avoiding perceptual pitfalls. J Acoust Soc Am 119:645–653. https://doi.org/10.1121/1.2139067

Deecke VB, Ford JKB, Spong P (1999) Quantifying complex patterns of bioacoustic variation: use of a neural network to compare killer whale (*Orcinus orca*) dialects. J Acoust Soc Am 105:2499–2507. https://doi.org/10.1121/1.426853

Delarue J, Todd SK, Van Parijs SM, Di Iorio L (2009) Geographic variation in Northwest Atlantic fin whale (*Balaenoptera physalus*) song: implications for stock structure assessment. J Acoust Soc Am 125:1774–1782. https://doi.org/10.1121/1.3068454

Delgado RA (2007) Geographic variation in the long sounds of male orangutans (*Pongo* spp.). Ethology 113:487–498. https://doi.org/10.1111/j.1439-0310.2007.01345.x

Deregnaucourt S, Guyomarch JC, Richard V (2001) Classification of hybrid crows in quail using artificial neural networks. Behav Process 56:103–112. https://doi.org/10.1016/S0376-6357(01)00188-7

Duda R, Hart P, Stork D (2001) Pattern classification, 2nd edn. Wiley, Hoboken, NJ

Dunlop RA, Noad MJ, Cato DH, Stokes D (2007) The social vocalization repertoire of east Australian migrating humpback whales (*Megaptera novaeangliae*). J Acoust Soc Am 122(5):2893–2905. https://doi.org/10.1121/1.2783115

Dunlop RA, Cato DH, Noad MJ, Stokes DM (2013) Source levels of social sounds in migrating humpback whales (*Megaptera novaeangliae*). J Acoust Soc Am 134(1):706–714. https://doi.org/10.1121/1.4807828

Egan JP (1975) Signal detection theory and ROC analysis. Academic Press, New York

Eiler KC, Banack SA (2004) Variability in the alarm call of golden-mantled ground squirrels (*Spermophilus lateralis* and *S. saturatus*). J Mammal 85:43–50. https://doi.org/10.1644/1545-1542(2004)085<0043:VITACO>2.0.CO;2

Erbe C, King AR (2008) Automatic detection of marine mammals using information entropy. J Acoust Soc Am 124(5):2833–2840. https://doi.org/10.1121/1.2982368

Erbe C, Verma A, McCauley R, Gavrilov A, Parnum I (2015) The marine soundscape of the Perth Canyon. Prog Oceanogr 137:38–51. https://doi.org/10.1016/j.pocean.2015.05.015

Erbe C, Reichmuth C, Cunningham K, Lucke K, Dooling R (2016) Communication masking in marine mammals: a review and research strategy. Mar Pollut Bull 103:15–38. https://doi.org/10.1016/j.marpolbul.2015.12.007

Erbe C, Dunlop R, Jenner KCS, Jenner M-NM, McCauley RD, Parnum I, Parsons M, Rogers T, Salgado-Kent C (2017) Review of underwater and in-air sounds emitted by Australian and Antarctic marine mammals. Acoust Aust 45:179–241. https://doi.org/10.1007/s40857-017-0101-z

Esfahanian M, Erdol N, Gerstein E, Zhuang H (2017) Two-stage detection of north Atlantic right whale upcalls using local binary patterns and machine learning algorithms. Appl Acoust 120:158–166. https://doi.org/10.1016/j.apacoust.2017.01.025

Fagerlund S (2007) Bird species recognition using support vector machines. EURASIP J Appl Sig Proc 2007(1):1–8. https://doi.org/10.1155/2007/38637

Fenton MB, Jacobson SL (1973) An automatic ultrasonic sensing system for monitoring the activity of some bats. Can J Zool 51:291–299. https://doi.org/10.1139/z73-041

Fitch WT (2003) Mammalian vocal production: themes and variation. In: Proceedings of the 1st International Conference on Acoustic Communication by Animals, 27–30 July, pp 81–82

Forti LR, Costa WP, Martins LB, Nunes-de-Almeida CH, Toledo LF (2016) Advertisement call and genetic structure conservatism: good news for an endangered Neotropical frog. PeerJ 4:e2014. https://doi.org/10.7717/peerj.2014

Freitag LE, Tyack PL (1993) Passive acoustic localization of the Atlantic bottlenose dolphin using whistles and echolocation clicks. J Acoust Soc Am 93:2197–2205. https://doi.org/10.1121/1.406681

Fristrup KM, Watkins WA (1993) Marine animal sound classification. Woods Hole Oceanographic Institution Technical Report WHOI-94-13, p 29

Frommolt K-H, Bardeli R, Clausen M (eds) (2007) Computational bioacoustics for assessing biodiversity. Proceed Internat Expert meeting on IT-based detection of bioacoustical patterns, 7–10 December 2007 at the International Academy for Nature Conservation (INA) Isle of Vilm, Germany. BfN - Skripten Federal Agency for Nature Conservation, p 234

Fukushima K, Wake N (1990) Alphanumeric character recognition by neocognitron. In: Miller RE (ed) Advanced neural computers. Elsevier Science, Amsterdam, pp 263–270

Fukuzawa Y, Webb WH, Pawley MD, Roper MM, Marsland S, Brunton DH, Gilman A (2020) Koe: web-based software to classify acoustic units and analyse sequence structure in animal vocalizations. Methods Ecol Evol 11:431–441. https://doi.org/10.1111/2041-210X.13336

Gannier A, Fuchs S, Quebre P, Oswald JN (2010) Performance of a contour-based classification method for whistles of Mediterranean dolphins. Appl Acoust 7:1063–1069. https://doi.org/10.1016/j.apacoust.2010.05.019

Gannon WL, Lawlor TE (1989) Variation in the chip vocalization of three species of Townsend's chipmunks (genus *Eutamias*). J Mammal 70:740–753

Gannon WL, Sherwin RE, deCarvalho TN, O'Farrell MJ (2001) Pinnae and echolocation call differences between *Myotis californicus* and *M. ciliolabrum* (Chiroptera: Vespertilionidae). Acta Chiropterol 3(1):77–91

Gannon WL, O'Farrell MJ, Corben C, Bedrick EJ (2004) Call character lexicon and analysis of field recorded bat echolocation calls. In: Thomas J, Moss C, Vater M (eds) Echolocation in bats and dolphins. The University of Chicago Press, Chicago, pp 478–484

Garland EC, Castellote M, Berchok CL (2015) Beluga whale (*Delphinapterus leucas*) vocalizations and call classification from the eastern Beaufort Sea population. J Acoust Soc Am 137:3054–3067. https://doi.org/10.1121/1.4919338

Garland EC, Rendell L, Lilley MS, Poole MM, Allen J, Noad MJ (2017) The devil is in the detail: quantifying vocal variation in a complex, multi-levelled, and rapidly evolving display. J Acoust Soc Am 142(1): 460–472. https://doi.org/10.1121/1.4991320

Gavrilov AN, Parsons MJG (2014) A MATLAB tool for the characterization of recorded underwater sound (CHORUS). Acoust Aust 42(3):190–196

Gavrilov A, McCauley R, Gedamke J (2012) Steady inter and intra-annual decrease in the vocalization frequency of Antarctic blue whales. J Acoust Soc Am 131(6): 4476–4480. https://doi.org/10.1121/1.4707425

Gedamke J, Costa DP, Dunstan A (2001) Localization and visual verification of a complex minke whale vocalization. J Acoust Soc Am 109(6):3038–3047. https://doi.org/10.1121/1.1371763

Gemello R, Mana F (1991) A neural approach to speaker independent isolated word recognition in an uncontrolled environment. In: Proceedings of the International Neural Networks Conference, Paris 9–13 July 1990, vol 1. Kluwer Academic Publishers, Dordrecht, pp 83–86

Ghosh J, Deuser LM, Beck SD (1992) A neural network based hybrid system for detection, characterization, and classification of short-duration oceanic signals. IEEE J Ocean Eng 17:351–363. https://doi.org/10.1109/48.180304

Gill SA, Bierema AM-K (2013) On the meaning of alarm calls: a review of functional reference in avian alarm calling. Ethology 119:449–461. https://doi.org/10.1111/eth.12097

Gillespie D, Caillat M (2008) Statistical classification of odontocete clicks. Can Acoust 36:20–26

Gillespie D, Caillat M, Gordon J (2013) Automatic detection and classification of odontocete whistles. J Acoust Soc Am 134:2427–2437. https://doi.org/10.1121/1.4816555

Gingras G, Fitch WT (2013) A three-parameter model for classifying anurans into four genera based on advertisement calls. J Acoust Soc Am 133:547–559. https://doi.org/10.1121/1.4768878

Goëau H, Glotin H, Vellinga WP, Planqué R, Joly A (2016) LifeCLEF bird identification task 2016: the arrival of deep learning. CLEF 1609:440–449

Griffin DR, Webster FA, Michael CR (1960) The echolocation of flying insects by bats. Anim Behav 8:141–154

Guemeur Y, Elisseeff A, Paugam-Moisey H (2000) A new multi-class SVM based on a uniform convergence result. Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium 4:183–188

Hamilton LJ, Cleary J (2010) Automatic discrimination of beaked whale clicks in noisy acoustic time series. In: OCEANS'10 IEEE Sydney, pp 1–5

Hammerschmidt K, Fischer J (1998) The vocal repertoire of Barbary macaques: a quantitative analysis of a graded signal system. Ethology 104(3):203–216. https://doi.org/10.1111/j.1439-0310.1998.tb00063.x

Hammerschmidt K, Reisinger E, Westekemper K, Ehrenreich L, Strenzke N, Fischer J (2012) Mice do not require auditory input for the normal development of their ultrasonic vocalizations. BMC Neurosci 13:40

Harland E (2008) Processing the workshop datasets using the TRUD algorithm. Can Acoust 36:27–33

He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. Proc IEEE Conf Comput Vis Pattern Recogn 2016:770–778

Helweg DA, Cato ADH, Jenkins PF, Garrigue D, McCauley RD (1998) Geographic variation in South Pacific humpback whale songs. Behaviour 135:1–27

Herr, A, Klomp, NL, Atkinson, JS (1997) Identification of bat echolocation calls using decision tree classification system Complexity International. https://www.researchgate.net/publication/293134471_Identification_of_bat_echolocation_calls_using_a_decision_tree_classification_system. Accessed 17 July 2017

Himawan I, Towsey M, Law B, Roe P (2018). Deep learning techniques for Koala Activity detection. In: INTERSPEECH, pp. 2107–2111

Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9(8):1735–1780

Holy TE, Guo Z (2005) Ultrasonic songs of male mice. PLoS One Biol 3(12):e386. https://doi.org/10.1371/journal.pbio.0030386

Horn AG, Falls JB (1996) Categorization and the design of signals: the case of song repertoires. In: Kroodsma DE, Miller EH (eds) Ecology and evolution of acoustic communication in birds. Comstock Publishing Associates, Ithaca, pp 121–135

Hotelling H (1933) Analysis of a complex of statistical variables into principal components. J Edu Psychol 24: 417–441

Huang X, Acero A, Hon H-W (2001) Spoken language processing. Prentice Hall, Upper Saddle River, NJ

Huang G, Liu Z, Van Der Maaten L, Weinberger KQ (2017) Densely connected convolutional networks. Proc IEEE Conf Comput Vis Pattern Recogn 2017: 4700–4708

Ibrahim AK, Chérubin LM, Zhuang H, Schärer Umpierre MT, Dalgleish F, Erdol N, Ouyang B, Dalgleish A (2018) An approach for automatic classification of grouper vocalizations with passive acoustic monitoring. J Acoust Soc Am 143:666–676. https://doi.org/10.1121/1.5022281

Itakura F (1975) Minimum prediction residual principle applied to speech recognition. IEEE Trans Acoust Speech Sig Process 23:57–72

Jacobson EK, Yack TM, Barlow J (2013) Evaluation of an automated acoustic beaked whale detection algorithm

using multiple validation and assessment methods. In: NOAA Technical Memorandum NOAA-TM-NMFS-SWFSC-509

Jaitly N, Hinton GE (2013) Vocal tract length perturbation (VTLP) improves speech recognition. In: Proceedings of ICML Workshop on Deep Learning for Audio, Speech and Language, vol 117

Janik VM (1999) Pitfalls in the categorization of behavior: a comparison of dolphin whistle classification methods. Anim Behav 57:133–143. https://doi.org/10.1006/anbe.1998.0923

Jarvis S, Dimarzio N, Morrissey R, Moretti D (2006) Automated classification of beaked whales and other small odontocetes in the Tongue of the Ocean, Bahamas. Oceans 2006:1–6. https://doi.org/10.1109/OCEANS.2006.307124

Jiang JJ, Bu LR, Duan FJ, Wang XQ, Liu W, Sun ZB, Li CY (2019) Whistle detection and classification for whales based on convolutional neural networks. Appl Acoust 150:169–178. https://doi.org/10.1016/j.apacoust.2019.02.007

Kandia V, Stylianou Y (2006) Detection of sperm whale clicks based on the Teager–Kaiser energy operator. Appl Acoust 67(11):1144–1163. https://doi.org/10.1016/j.apacoust.2006.05.007

Karlsen JD, Bisther A, Lyndersen C, Haug T, Kovacs KM (2002) Summer vocalizations of adult male white whales (*Delphinapterus leucas*) in Svalbard, Norway. Polar Biol 25:808–817. https://doi.org/10.1007/s00300-002-0415-6

Keen S, Ross JC, Griffiths ET, Lanzone M, Farnsworth A (2014) A comparison of similarity-based approaches in the classification of flight calls of four species of North American wood-warblers (Parulidae). Ecol Inf 21:25–33. https://doi.org/10.1016/j.ecoinf.2014.01.001

Keighley MV, Langmore NE, Zdenek CN, Heinsohn R (2017) Geographic variation in the vocalizations of Australian palm cockatoos (*Probosciger aterrimus*). Bioacoustics 26(1):91–108. https://doi.org/10.1080/09524622.2016.1201778

Kershenbaum A, Blumstein DT, Roch MA, Akcay C, Backus G, Bee MA, Bohn K, Cao Y, Carter G, Cäsar C, Coen M, DeRuiter SL, Doyle L, Edelman S, Ferrer-i-Cancho R, Freeberg TM, Garland EC, Gustison M, Harley HE, Huetz C, Hughes M, Bruno JH, Ilany A, Jin DZ, Johnson M, Ju C, Karnowski J, Lohr B, Manser MB, McCowan B, Mercado E, Narins PM, Piel A, Rice M, Salmi R, Sasahara K, Sayigh L, Shiu Y, Taylor C, Vallejo EE, Waller S, Zamora-Gutierrez V (2016) Acoustic sequences in non-human animals: a tutorial review and prospectus. Biol Rev 91:13–52

Kingma DP, Welling M (2013) Auto-encoding variational bayes. arXiv preprint arXiv:1312.6114

Klinck H, Mellinger DK (2011) The energy ratio mapping algorithm: a tool to improve the energy-based detection of odontocete echolocation clicks. J Acoust Soc Am 129(4):1807–1812. https://doi.org/10.1121/1.3531924

Ko T, Peddinti V, Povey D, Khudanpur S (2015) Audio augmentation for speech recognition. In: Sixteenth Annual Conference of the International Speech Communication Association

Kogan J, Margoliash D (1998) Automated recognition of bird song elements from continuous recordings using dynamic time warping and hidden Markov models: a comparative study. J Acoust Soc Am 103:2185–2196. https://doi.org/10.1121/1.421364

Kollmorgen S, Hahnloser RH, Mante V (2020) Nearest neighbours reveal fast and slow components of motor learning. Nature 577(7791):526–530. https://doi.org/10.1038/s41586-019-1892-x

Kondo N, Watanabe S (2009) Contact calls: information and social function. Jpn Psych Res 51:197–208. https://doi.org/10.1111/j.1468-5884.2009.00399.x

Koren L, Geffen E (2009) Complex call in male rock hyrax (*Procavia capensis*): a multi-information distributing channel. Behav Ecol Sociobiol 63(4):581–590. https://doi.org/10.1007/s00265-008-0693-2

Koren L, Geffen E (2011) Individual identity is communicated through multiple pathways in male rock hyrax (*Procavia capensis*) songs. Behav Ecol Sociobiol 65(4):675–684. https://doi.org/10.1007/s00265-010-1069-y

Koren L, Mokady O, Geffen E (2008) Social status and cortisol levels in singing rock hyraxes. Horm Behav 54:212–216

Krizhevsky A, Sutskever I, Hinton GE (2017) ImageNet classification with deep convolutional neural networks. Commun ACM 60(6):84–90

Kruskal J, Sankoff D (1983) An anthology of algorithms and concepts for sequence comparison. In: Sankoff D, Kruskal J (eds) Time warps, string edits and macromolecules: the theory and practice of string comparison. Addison-Wesley, Reading, MA, pp 265–310

Lammers MO, Au WWL, Herzing DL (2003) The broadband social acoustic signaling behavior of spinner and spotted dolphins. J Acoust Soc Am 114:1629–1639. https://doi.org/10.1121/1.1596173

Law BS, Reinhold L, Pennay M (2002) Geographic variation in the echolocation sounds of *Vespadelus* spp. (Vespertilionidae) from New South Wales and Queensland, Australia. Acta Chiropt 4:201–215. https://doi.org/10.3161/001.004.0208

Le Boeuf BJ, Peterson RS (1969) Dialects in elephant seals. Science 166(3913):1654–1656. https://doi.org/10.1126/science.166.3913.1654

Leblanc E, Bahoura M, Simard Y (2008) Comparison of automatic classification methods for beluga whale vocalizations. J Acoust Soc Am 123:3772

LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, Jackel LD (1989a) Backpropagation applied to handwritten zip code recognition. Neural Comput 1(4):541–551. https://doi.org/10.1162/neco.1989.1.4.541

LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, Jackel LD (1989b) Handwritten digit recognition with a back-propagation network. In:

Proceedings of the 2nd International Conference on Neural Information Processing Systems, pp 396–404
LeCun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. Proc IEEE 86(11):2278–2324. https://doi.org/10.1109/5.726791
LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436–444. https://doi.org/10.1038/nature14539
Lee C-H, Hsu S-B, Shih J-L, Chou C-H (2013) Continuous birdsong recognition using Gaussian mixture modeling of image shape features. IEEE Trans Multimedia 15:454–464. https://doi.org/10.1109/TMM.2012.2229969
Leonard ML, Horn AG (2001) Begging calls and parental feeding decisions in tree swallows (*Tachycineta bicolor*). Behav Ecol Sociobiol 49:170–175. https://doi.org/10.1007/s002650000290
Levinson S (1985) Structural methods in automatic speech recognition. Proc IEEE 73:1625–1648. https://doi.org/10.1109/PROC.1985.13344
Li Z, Tang S, Yan S (2002) Multi-class SVM classifier based on pair wise coupling. In: Proceedings of the First International Workshop, SVM 2002, Niagara Falls, Canada, p 321
Liaw A, Wiener M (2002) Classification and regression by Random Forest. R News 2:18–22
Linderman GC, Rachh M, Hoskins JG, Steinerberger S, Kluger Y (2017) Efficient algorithms for t-distributed stochastic neighborhood embedding. arXiv preprint arXiv:1712.09005
Lippman R (1989) Pattern classification using neural networks. IEEE Commun Mag 1989:47–64
Luo W, Yang W, Zhang Y (2019) Convolutional neural network for detecting odontocete echolocation clicks. J Acoust Soc Am 145(1):EL7–EL12. https://doi.org/10.1121/1.5085647
Maaten LV (2014) Accelerating t-SNE using tree-based algorithms. J Mach Learn Res 15(1):3221–3245
Maaten LV, Hinton G (2008) Visualizing data using t-SNE. J Mach Learn Res 9:2579–2605
Mac Aodha O, Gibb R, Barlow KE, Browning E, Firman M, Freeman R, Harder B, Kinsey L, Mead GR, Newson SE, Pandourski I (2018) Bat detective—deep learning tools for bat acoustic signal detection. PLoS Comput Biol 14(3):e1005995. https://doi.org/10.1371/journal.pcbi.1005995
Madhusudhana S, Gavrilov AN, Erbe C (2015) Automatic detection of echolocation clicks based on a Gabor model of their waveform. J Acoust Soc Am 137(6):3077–3086. https://doi.org/10.1121/1.4921609
Madhusudhana S, Symes LB, Klinck H (2019) A deep convolutional neural network based classifier for passive acoustic monitoring of neotropical katydids. J Acoust Soc Am 146(4):2982–2982. https://doi.org/10.1121/1.5137323
Madhusudhana S, Murray A, Erbe C (2020) Automatic detectors for low-frequency vocalizations of Omura's whales, *Balaenoptera omurai*: a performance comparison. J Acoust Soc Am 147(5):3078–3090. https://doi.org/10.1121/10.0001108
Madhusudhana S, Shiu Y, Klinck H, Fleishman E, Liu X, Nosal EM, Helble T, Cholewiak D, Gillespie D, Širović A, Roch MA (2021) Improve automatic detection of animal call sequences with temporal context. J R Soc Interface 18:20210297. https://doi.org/10.1098/rsif.2021.0297
Madsen PT, Surlykke A (2013) Functional convergence in bat and toothed whale biosonars. Physiology 28(5):276–283. https://doi.org/10.1152/physiol.00008.2013
Makhoul J, Schwarz R (1995) State of the art in continuous speech recognition. Proc Nat Acad Sci USA 92:9956–9963. https://doi.org/10.1073/pnas.92.22.9956
Malfante M, Mohammed O, Gervaise C, Dalla Mura M, Mars JI (2018) Use of deep features for the automatic classification of fish sounds. In: 2018 OCEANS-MTS/IEEE Kobe Techno-Oceans (OTO), pp 1–5. https://doi.org/10.1109/OCEANSKOBE.2018.8559276
Mankin RW, Smith T, Tropp JM, Atkinson EB, Young DY (2008) Detection of *Anoplophora glabripennis* (Coleoptera: Cerambycidae) larvae in different host trees and tissues by automated analysis of sound-impulse frequency and temporal patterns. J Econ Entomol 101(3):838–849. https://doi.org/10.1093/jee/101.3.838
Marler P (2004) Bird calls: a cornucopia for communication. In: Marler P, Slabbekoorn H (eds) Nature's music: the science of birdsong. Elsevier, Amsterdam, pp 132–177
Martindale S (1980a) On the multivariate analysis of avian vocalizations. J Theor Biol 83:107–110. https://doi.org/10.1016/0022-5193(80)90374-4
Martindale S (1980b) A numeric approach to the analysis of solitary vireo songs. Condor 82:199–211. https://doi.org/10.2307/1367478
Mazhar S, Ura T, Bahl R (2007) Vocalization based individual classification of humpback whales using support-vector-machine. Oceans 2007:1–9. https://doi.org/10.1109/OCEANS.2007.4449356
McDonald MA, Mesnick SL, Hildebrand JA (2006) Biogeographic characterisation of blue whale song worldwide: using song to identify populations. J Cetacean Res Manag 8(1):55–65
McInnes L, Healy J, Melville J (2018) UMAP: uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:1802.03426
McLaughlin J, Josso N, Ioana C (2008) Detection and classification of sound types in the vocalizations of north-east pacific blue whales. J Acoust Soc Am 123:3102
McLister D, Stevens ED, Bogart JP (1995) Comparative contractile dynamics of calling and locomotor muscles in three hylid frogs. J Exp Biol 198(7):1527–1538. https://doi.org/10.1242/jeb.198.7.1527
Mellinger DK (2008) A neural network for classifying clicks of Blainville's beaked whales (*Mesoplodon densirostris*). Can Acoust 36:55–59
Mellinger DK, Bradbury JW (2007) Acoustic measurement of marine mammal sounds in noisy

environments. In: Proceedings of the 2nd International Conference on Underwater Acoustic Measurements: Technologies and Results, Heraklion, Greece, 25–29 June 2007

Mellinger DK, Clark CW (2000) Recognizing transient low-frequency whale sounds by spectrogram correlation. J Acoust Soc Am 107(6):3518–3529. https://doi.org/10.1121/1.429434

Mellinger DK, Martin SW, Morrissey RP, Thomas L, Yosco JJ (2011) A method for detecting whistles, moans and other frequency contour sounds. J Acoust Soc Am 129:4055–4061. https://doi.org/10.1121/1.3531926

Mendelson TC, Shaw KL (2003) Rapid speciation in an arthropod. Nature 433:375–376. https://doi.org/10.1038/433375a

Mitani JC, Hasegawa T, Groslouis J, Marler P, Byrne R (1992) Dialects in wild chimpanzees. Am J Primatol 27:233–243

Møhl B, Wahlberg M, Madsen PT, Heerford A, Lund A (2003) The monopulsed nature of sperm whale sonar clicks. J Acoust Soc Am 114(2):1143–1154. https://doi.org/10.1121/1.1586258

Moon TK (1996) The expectation-maximization algorithm. IEEE Sig Process Mag 13:47–60. https://doi.org/10.1109/79.543975

Morrissey RP, Ward J, DiMarzio N, Jarvis S, Moretti DJ (2006) Passive acoustic detection and localization of sperm whales (*Physeter macrocephalus*) in the tongue of the ocean. Appl Acoust 67:1091–1105. https://doi.org/10.1016/j.apacoust.2006.05.014

Mouy X, Leary D, Martin B, Laurinolli M (2008) A comparison of methods for the automatic classification of marine mammal vocalizations in the Arctic. In: Proceedings of the PASSIVE'08 Workshop on New Trends for Environmental Monitoring using Passive Systems, Hyeres, France, 14–17 October 2008

Murray SO, Mercado E, Roitblat HL (1998) Characterizing the graded structure of false killer whale (*Pseudorca crassidens*) vocalizations. J Acoust Soc Am 104:1679–1687. https://doi.org/10.1121/1.424380

Myers C, Rabiner LR, Rosenberg AE (1980) Performance tradeoffs in dynamic time warping algorithms for isolated word recognition. IEEE Trans Acoust Speech Sig Process 28:623–635. https://doi.org/10.1109/TASSP.1980.1163491

Nagy CM, Rockwell RF (2012) Identification of individual eastern screech-owls (*Megascops asio)* via vocalization analysis. Bioacoustics 21:127–140. https://doi.org/10.1080/09524622.2011.651829

Narins PM, Feng AS, Fay RR (eds) (2006) Hearing and sound communication in amphibians. Springer, New York

Noad MJ, Cato DH, Bryden MM, Jenner MN, Jenner KCS (2000) Cultural revolution in whale songs. Nature 408:537. https://doi.org/10.1038/35046199

Noda JJ, Travieso CM, Sánchez-Rodríguez D (2016) Automatic taxonomic classification of fish based on their acoustic signals. Appl Sci 6(12):443. https://doi.org/10.3390/app6120443

O'Farrell MJ, Miller BW, Gannon WL (1999) Qualitative identification of free-flying bats using Anabat detector. J Mammal 80:11–23

Oh J, Laubach M, Luczak A (2003) Estimating neuronal variable importance with random forest. Proc IEEE Bioeng Conf:33–34. https://doi.org/10.1109/NEBC.2003.1215978

Oleson EM, Širović A, Bayless AR, Hildebrand JA (2014) Synchronous seasonal change in fin whale song in the North Pacific. PLoS One 9(12):e115678. https://doi.org/10.1371/journal.pone.0115678

Oswald JN, Barlow J, Norris TF (2003) Acoustic identification of nine delphinid species in the eastern tropical Pacific Ocean. Mar Mamm Sci 19:20–37. https://doi.org/10.1111/j.1748-7692.2003.tb01090.x

Oswald JN, Rankin S, Barlow J, Lammers MO (2007) A tool for real-time acoustic species identification of delphinid whistles. J Acoust Soc Am 122:587–595. https://doi.org/10.1121/1.2743157

Oswald JN, Au WWL, Duennebier F (2011) Minke whale (*Balaenoptera acutorostrata*) boings detected at the Station ALOHA cabled observatory. J Acoust Soc Am 129:3353–3360. https://doi.org/10.1121/1.3575555

Oswald JN, Rankin S, Barlow J, Oswald M (2013) Real-time odontocete call classification algorithm: software for species identification of delphinid whistles. In: Adam O, Samaran F (eds) Detection, classification and localization of marine mammals using passive acoustics, 2003-2013: 10 years of international research. DIRAC NGO, Paris, France

Oswald JN, Walmsley SF, Casey C, Fregosi S, Southall B, Janik VM (2021) Species information in whistle frequency modulation patterns of common dolphins. Philos Trans R Soc B 376:20210046. https://doi.org/10.1098/rstb.2021.0046

Ou H, Au WWL, Oswald JN (2012) A non-spectrogram-correlation method of automatically detecting minke whale boings. J Acoust Soc Am 132:EL317–EL322

Ouattara K, Lemasson A, Zuberbunter K (2009) Campbell's monkeys concatenate vocalizations into context-specific call sequences. Proc Natl Acad Sci USA 106(51):22026

Papale E, Azzolin M, Cascao I, Gannier A, Lammers MO, Martin VM, Oswald JN, Perez-Gil M, Prieto R, Silva MA, Giacoma C (2013) Geographic variability in the acoustic parameters of striped dolphin's (*Stenella coeruleoalba*) whistles. J Acoust Soc Am 133:1126–1134. https://doi.org/10.1121/1.4774274

Papale E, Azzolin M, Cascao I, Gannier A, Lammers MO, Martin VM, Oswald J, Perez-Gil M, Prieto R, Silva MA, Giacoma C (2014) Macro- and micro- geographic variation of short-beaked common dolphin's whistles in the Mediterranean Sea and Atlantic Ocean. Ethol Ecol Evol 26:392–404. https://doi.org/10.1080/03949370.2013.851122

Park DS, Chan W, Zhang Y, Chiu C, Zoph B, Cubuk ED, Le QV (2019) SpecAugment: a simple data

augmentation method for automatic speech recognition. Proc Interspeech 2019:2613–2617. https://doi.org/10.21437/Interspeech.2019-2680

Parsons S, Boonman AM, Obrist MK (2000) Advantages and disadvantages of techniques for transforming and analyzing chiropteran echolocation calls. J Mammal 81:927–938. https://doi.org/10.1644/1545-1542(2000)081<0927:AADOTF>2.0.CO;2

Payne K, Payne R (1985) Large scale changes over 19 years in songs of humpback whales in Bermuda. Z Tierpsychol 68:89–114. https://doi.org/10.1111/j.1439-0310.1985.tb00118.x

Picone JW (1993) Signal modeling techniques in speech recognition. Proc IEEE 81:1215–1247. https://doi.org/10.1109/5.237532

Placer J, Slobodchikoff CN (2000) A fuzzy-neural system for identification of species-specific alarm sounds of Gunnison's prairie dogs. Behav Process 52:1–9. https://doi.org/10.1016/S0376-6357(00)00105-4

Potter JR, Mellinger DK, Clark CW (1994) Marine mammal sound discrimination using artificial neural networks. J Acoust Soc Am 96:1255–1262. https://doi.org/10.1121/1.410274

Pozzi L, Gamba M, Giacoma C (2010) The use of Artificial Neural Networks to classify primate vocalizations: a pilot study on black lemurs. Am J Primatol 72(4):337–348. https://doi.org/10.1002/ajp.20786

Pröhl H, Koshy RA, Mueller U, Rand AS, Ryan MJ (2006) Geographic variation of genetic and behavioral traits in northern and southern Túngara frogs. Evol 60:1669–1679. https://doi.org/10.1111/j.0014-3820.2006.tb00511.x

Rabiner LR (1989) A tutorial on hidden Markov models and selected applications in speech recognition. Proc IEEE 77:257–285

Rabiner LR, Juang BH (1986) An introduction to Hidden Markov Models. IEEE ASSP Mag 1986:4–16

Rabiner LR, Levinson S, Sondhi M (1983) On the application of vector quantization and hidden Markov models to speaker-independent, isolated word recognition. Bell Syst Tech J 62:1075–1106. https://doi.org/10.1002/j.1538-7305.1983.tb03115.x

Rabiner LR, Juang B, Lee C (1996) An overview of automatic speech recognition. In: Lee C, Soong F, Paliwal K (eds) Automatic speech and speaker recognition. Kluwer Academic, New York, pp 1–30

Rankin S, Barlow J (2005) Source of the North Pacific 'boing' sound attributed to minke whales. J Acoust Soc Am 118(5):3346–3351. https://doi.org/10.1121/1.2046747

Rankin S, Ljungblad D, Clark CW, Kato H (2005) Vocalisations of Antarctic blue whales, *Balaenoptera musculus intermedia*, recorded during the 2001/2002 and 2002/2003 IWC/SOWER circumpolar cruises, Area V, Antarctica. J Cet Res Manag 7(1):13–20

Rankin S, Archer F, Keating JL, Oswald JN, Oswald M, Curtis A, Barlow J (2016) Acoustic classification of dolphins in the California Current using whistles, clicks and burst-pulses. Mar Mamm Sci 33:520–540. https://doi.org/10.1111/mms.12381

Reby D, André-Obrecht R, Galinier A, Farinas J, Cargnelutti B (2006) Cepstral coefficients and hidden Markov models reveal idiosyncratic voice characteristics in red deer (*Cervus elaphus*) stags. J Acoust Soc Am 120:4080–4089. https://doi.org/10.1121/1.2358006

Recalde-Salas A, Salgado Kent CP, Parsons MJG, Marley SA, McCauley RD (2014) Non-song vocalizations of pygmy blue whales in Geographe Bay, Western Australia. J Acoust Soc Am 135(5):EL213–EL218. https://doi.org/10.1121/1.4871581

Recalde-Salas A, Erbe C, Salgado Kent C, Parsons M (2020) Non-song vocalizations of humpback whales in Western Australia. Front Mar Sci 7:141. https://doi.org/10.3389/fmars.2020.00141

Rickwood P, Taylor A (2008) Methods for automatically analyzing humpback song units. J Acoust Soc Am 123:1763–1772. https://doi.org/10.1121/1.2836748

Risch D, Gales NJ, Gedamke J, Kindermann L, Nowacek DP, Read AJ, Siebert U, Van Opzeeland IC, Van Parijs SM, Friedlander AS (2014) Mysterious bio-duck sound attributed to the Antarctic minke whale (*Balaenoptera bonaerensis*). Biol Lett 10:20140175. https://doi.org/10.1098/rsbl.2014.0175

Roch MA, Soldevilla MS, Burtenshaw JC, Henderson EE, Hildebrand JA (2007) Gaussian mixture model classification of odontocetes in the Southern California Bight and the Gulf of California. J Acoust Soc Am 121:1737–1748. https://doi.org/10.1121/1.2400663

Roch MA, Soldevilla MS, Hoenigman R, Wiggins SM, Hildebrand JA (2008) Comparison of machine-learning techniques for the classification of echolocation clicks from three species of odontocetes. Can Acoust 36:41–47

Roch MA, Brandes TS, Patel B, Barkley Y, Baumann-Pickering S, Soldevilla MS (2011) Automated extraction of odontocete whistle contours. J Acoust Soc Am 130:2212–2223. https://doi.org/10.1121/1.3624821

Rocha HS, Ferreira LS, Paula BC, Rodrigues HG, Sousa-Lima RS (2015) An evaluation of manual and automated methods for detecting sounds of mane wolves (*Chrysocyon brachyurus* Illiger 1815). Bioacoustics 24:185–198. https://doi.org/10.1080/09524622.2015.1019361

Roitblat HL, Moore PWB, Nachtigall PE, Penner RH, Au WWL (1989) Natural echolocation with an artificial neural network. Int J Neural Syst 1:239–247

Rosenblatt F (1958) The perceptron: a probabilistic model for information storage and organization in the brain. Psychol Rev 65:386–408. https://doi.org/10.1037/h0042519

Ross JC, Allen PE (2014) Random forest for improved analysis efficiency in passive acoustic monitoring. Ecol Inform 21:34–39. https://doi.org/10.1016/j.ecoinf.2013.12.002

Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagating errors. Nature 323(6088):533–536. https://doi.org/10.1038/323533a0

Russo D, Mucedda M, Bello M, Biscardi S, Pidinchedda E, Jones G (2007) Divergent echolocation sound frequencies in insular rhinolophids (Chiroptera): a case of character displacement? J Bioeng 34:2129–2138. https://doi.org/10.1111/j.1365-2699.2007.01762.x

Sainburg T, Theilman B, Thielk M, Gentner TQ (2019) Parallels in the sequential organization of birdsong and human speech. Nat Commun 10:3636. https://doi.org/10.1038/s41467-019-11605-y

Sakoe H, Chiba S (1978) Dynamic programming optimization for spoken word recognition. IEEE Trans Acoust Speech Sig Process 26:43–49. https://doi.org/10.1109/TASSP.1978.1163055

Schassburger RM (1993) Vocal communication in the timber wolf, *Canis lupus*, Linnaeus: structure, motivation, and ontogeny. Parey Scientific Publication, New York

Schon PC, Puppe B, Manteauffel G (2001) Linear prediction coding analysis and self-organizing feature map as tools to classify stress sounds of domestic pigs (*Sus scrofa*). J Acoust Soc Am 110:1425–1431. https://doi.org/10.1121/1.1388003

Sethi SS, Jones NS, Fulcher BD, Picinali L, Clink DJ, Klinck H, Orme CD, Wrege PH, Ewers RM (2020) Characterizing soundscapes across diverse ecosystems using a universal acoustic feature set. Proc Natl Acad Sci 117(29):17049–17055. https://doi.org/10.1073/pnas.2004702117

Shannon CE, Weaver W (1998) The mathematical theory of communication. University of Illinois Press, Champaign

Shiu Y, Palmer KJ, Roch MA, Fleishman E, Liu X, Nosal EM, Helble T, Cholewiak D, Gillespie D, Klinck H (2020) Deep neural networks for automated detection of marine mammal species. Sci Rep 10(1):1–12. https://doi.org/10.1038/s41598-020-57549-y

Sibley DA (2000) The Sibley field guide to birds. Knopf, New York

Simmons JA, Wever EG, Pylka JM (1971) Periodical cicada: sound production and hearing. Science 171(3967):212–213. https://doi.org/10.1126/science.171.3967.212

Širović A (2016) Variability in the performance of the spectrogram correlation detector for north-east Pacific blue whale calls. Bioacoustics 25(2):145–160. https://doi.org/10.1080/09524622.2015.1124248

Širović A, Cutter GR, Butler JL, Demer DA (2009) Rockfish sounds and their potential use for population monitoring in the Southern California Bight. ICES J Mar Sci 66:981–990. https://doi.org/10.1093/icesjms/fsp064

Sjare B, Stirling I, Spencer C (2003) Seasonal and longer-term variability in the songs of Atlantic walruses breeding in the Canadian High Arctic. Aquat Mamm 29(2):297–318

Slobodchikoff CN, Ackers SH, Van Ert M (1998) Geographic variation in alarm calls of Gunnison's prairie dogs. J Mammal 79(4):1265–1272. https://doi.org/10.2307/1383018

Somervuo P, Härmä A, Fagerlund S (2006) Parametric representations of bird sounds for automatic species recognition. IEEE Trans Audio Speech Lang Process 14:2252–2263. https://doi.org/10.1109/TASL.2006.872624

Sparling DW, Williams JD (1978) Multivariate analysis of avian vocalizations. J Theor Biol 74:83–107. https://doi.org/10.1016/0022-5193(78)90291-6

Stafford KM, Fox CG, Clark DS (1998) Long-range acoustic detection and localization of blue whale sounds in the northeast Pacific Ocean. J Acoust Soc Am 104(6):3616–3625. https://doi.org/10.1121/1.423944

Stafford KM, Nieukirk SL, Fox CG (1999) Low-frequency whale sounds recorded on hydrophones moored in the eastern tropical Pacific. J Acoust Soc Am 106:3687–3698. https://doi.org/10.1121/1.428220

Stafford KM, Moore SE, Laidre KL, Heide-Jørgensen MP (2008) Bowhead whale springtime song off West Greenland. J Acoust Soc Am 124(5):3315–3323. https://doi.org/10.1121/1.2980443

Starnberger I, Preininger D, Hödl W (2014) The anuran vocal sac: a tool for multimodal signalling. Anim Behav 97:281–288. https://doi.org/10.1016/j.anbehav.2014.07.027

Stoeger AS, Heilmann G, Zeppelzauer M, Ganswindt A, Hensman S, Charlton BD (2012) Visualizing sound emission of elephant vocalizations: evidence for two rumble production types. PLoS One 7:1–8. https://doi.org/10.1371/journal.pone.0048907

Stowell D, Wood M, Stylianou Y, Glotin H (2016). Bird detection in audio: a survey and a challenge. In: 2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP), pp 1–6. https://doi.org/10.1109/MLSP.2016.7738875

Sturtivant C, Datta S (1997) Automatic dolphin whistle detection, extraction, encoding, and classification. Proc Inst Acoust 19:259–266

Suzuki R, Buck J, Tyack P (2006) Information entropy of humpback whale songs. J Acoust Soc Am 119:1849–1866. https://doi.org/10.1121/1.2161827

Swets JA, Dawes RM, Monahan J (2000) Better decisions through science. Sci Am 283:82–87

Takahashi N, Kashino M, Hironaka N (2010) Structure of rat ultrasonic vocalizations and its relevance to behavior. PLoS One 5(11):e14115. https://doi.org/10.1371/journal.pone.0014115

Tan M, McDonald K (2017) Bird sounds | Experiments with Google [online]. https://experiments.withgoogle.com/bird-sounds

Tchernichovski O, Nottebohm F, Ho CE, Pesaran B, Mitra PP (2000) A procedure for an automated measurement of song similarity. Anim Behav 59:1167–1176. https://doi.org/10.1006/anbe.1999.1416

Tenenbaum JB, De Silva V, Langford JC (2000) A global geometric framework for nonlinear dimensionality reduction. Science 290(5500):2319–2323. https://doi.org/10.1126/science.290.5500.2319

Thomas JA, Golladay CL (1995) Analysis of underwater vocalizations of leopard seals (*Hydrurga leptonyx*). In: Kastelein RA, Thomas JA, Nachtigall PE (eds) Sensory systems of aquatic mammals. De Spil Publishers, Amsterdam, pp 201–221

Thomas M, Martin B, Kowarski K, Gaudet B, Matwin S (2019) Marine mammal species classification using convolutional neural networks and a novel acoustic representation. In: Joint European Conference on Machine Learning and Knowledge Discovery in Databases, pp 290–305

Torrey L, Shavlik J (2010) Transfer learning. In: Handbook of research on machine learning applications and trends: algorithms, methods, and techniques. IGI Global, New York, pp 242–264

Trawicki MB, Johnson MT, Osiejuk TS (2005) Automatic song-type classification and speaker identification of Norwegian ortolan bunting. IEEE Int Conf Mach Learn Sig Process (MLSP) 2005:277–282. https://doi.org/10.1109/MLSP.2005.1532913

Trifa VM, Kirschel ANG, Taylor CE (2008) Automated species recognition of antbirds in a Mexican rainforest using hidden Markov Models. J Acoust Soc Am 123:2424–2431. https://doi.org/10.1121/1.2839017

Valente D, Wang H, Andrews P, Mitra PP, Saar S, Tchernichovski O, Golani I, Benjamini Y (2007) Characterizing animal behavior through audio and video signal processing. IEEE Multimedia 14:32–41. https://doi.org/10.1109/MMUL.2007.71

Van Allen E, Menon MM, Dicaprio N (1990) A modular architecture for object recognition using neural networks. In: Proceedings of International Neural Networks Conference, Paris, vol 1, pp 35–379, 13 July 1990. Kluwer Academic Publishers, Dordrecht

Vapnik VN (1998) Statistical learning theory. Wiley, New York

Venter PJ, Hanekom JJ (2010) Automatic detection of African elephant (*Loxodonta africana*) infrasonic vocalizations from recordings. Biosyst Eng 106:286–294. https://doi.org/10.1016/j.biosystemseng.2010.04.001

Von Muggenthaler E, Reinhart P, Lympany B, Craft RB (2003) Songlike vocalizations from the Sumatran rhinoceros (*Dicerorhinus sumatrensis*). Acoust Res Lett 4(3):83–88. https://doi.org/10.1121/1.1588271

Waibel A, Hanazawa T, Hinton G, Shikano K, Lang KL (1989) Phoneme recognition using time-delay neural networks. IEEE Trans Acoust Speech Signal Proc 37:328–339. https://doi.org/10.1109/29.21701

Ward J, Morrissey R, Moretti D, DiMarzio N, Jarvis S, Johnson M, Tyack PL, White C (2008) Passive acoustic detection and localization of *Mesoplodon densirostris* (Blainville's beaked whale) vocalizations using distributed bottom-mounted hydrophones in conjunction with a digital tag (DTag) recording. Can Acoust 36:60–66

Ward R, Parnum I, Erbe C, Salgado-Kent CP (2016) Whistle characteristics of Indo-Pacific bottlenose dolphins (*Tursiops aduncus*) in the Fremantle Inner Harbour, Western Australia. Acoust Aust 44(1):159–169. https://doi.org/10.1007/s40857-015-0041-4

Ward R, Gavrilov AN, McCauley RD (2017) "Spot" call: A common sound from an unidentified great whale in Australian temperate waters. J Acoust Soc Am 142(2):EL231–EL236. https://doi.org/10.1121/1.4998608

Weisburn BA, Mitchell SG, Clark CW, Parks TW (1993) Isolating biological acoustic transient signals. Proc IEEE Int Conf Acoust Speech Sig Process 1:269–272. https://doi.org/10.1109/ICASSP.1993.319107

Wellard R, Erbe C, Fouda L, Blewitt M (2015) Vocalisations of killer whales (*Orcinus orca*) in the Bremer Canyon, Western Australia. PLoS One 10(9):e0136535. https://doi.org/10.1371/journal.pone.0136535

Wells KD (2007) The ecology and behaviour of amphibians. University of Chicago Press, Chicago, IL

Wich SA, Schel AM, De Vries H (2008) Geographic variation in Thomas langur (*Presbytis thomasi*) loud sounds. Am J Primatol 70:566–574. https://doi.org/10.1002/ajp.20527

Winn HE, Winn LK (1978) The song of the humpback whale *Megaptera novaeangliae* in the West Indies. Mar Biol 47:97–114. https://doi.org/10.1007/BF00395631

Wood JD, McCowan B, Langbauer WR, Viljoen JJ, Hart LA (2005) Classification of African elephant *Loxodonta africana* rumbles using acoustic parameters and cluster analysis. Bioacoustics 15:143–161. https://doi.org/10.1080/09524622.2005.9753544

Yamamoto O, Moore B, Brand L (2001) Variation in the bark sound of the red squirrel (*Tamiasciurus hudsonicus*). West N Am Nat 61:395–402

Yang X-J, Lei F-M, Wang G, Jesse AJ (2007) Syllable sharing and inter-individual syllable variation in Anna's hummingbird *Calypte anna* songs, in San Francisco, California. Folia Zool 56:307–318

Yoshino H, Armstrong KN, Izawa M, Yokoyama J, Kawata M (2008) Genetic and acoustic population structuring in the Okinawa least horseshoe bat: are intercolony acoustic differences maintained by vertical maternal transmission? Mol Ecol 17:4978–4991. https://doi.org/10.1111/j.1365-294X.2008.03975.x

Zar JH (2009) Biostatistical analysis, 5th edn. Pearson, New York, p 960

Zeppelzauer M, Hensman S, Stoeger AS (2015) Towards an automated acoustic detection system for free-ranging elephants. Bioacoustics 24:13–29. https://doi.org/10.1080/09524622.2014.906321

Zhang YJ, Huang JF, Gong N, Ling ZH, Hu Y (2018) Automatic detection and classification of marmoset vocalizations using deep and recurrent neural

networks. J Acoust Soc Am 144(1):478–487. https://doi.org/10.1121/1.5047743

Zhong M, LeBien J, Campos-Cerqueira M, Dodhia R, Ferres JL, Velev JP, Aide TM (2020) Multispecies bioacoustic classification using transfer learning of deep convolutional neural networks with pseudo-labeling. Appl Acoust 166:107375. https://doi.org/10.1016/j.apacoust.2020.107375

Zuberbuhler K, Jenny D, Bshary R (1999) The predator deterrence function of primate alarm calls. Ethology 105:477–490. https://doi.org/10.1046/j.1439-0310.1999.00396.x

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



# Fundamental Data Analysis Tools and Concepts for Bioacoustical Research

9

Chandra Salgado Kent, Tiago A. Marques, and Danielle Harris

## 9.1 Introduction

Bioacoustics has emerged as a prominent, non-invasive, and innovative approach to obtaining scientific knowledge about animal behavior and ecology. As a consequence, bioacousticians play an important role in today's societies, often informing decision-makers in governments, industries, and communities. As an example, bioacousticians are often asked whether a species, a population, a community, or individual animals will sustain impacts from noise—or any other impact, of course, but noise is particularly relevant to the running theme of the book—generated from particular human activities. Sometimes, government regulators require "yes" or "no" answers to these questions. A knowledgeable bioacoustician, any scientist in fact, will know that usually it is difficult to provide simple 'yes' or 'no' answers. This is because the magnitude of impact that is biologically significant is usually not known. For instance, imagine the question relates to whether loud construction works will result in a decline of a local population of animals. The observed impact is that animals reduce the time spent feeding. Therefore, the required reduction in time feeding that will lead to a population decline must be known to be able to provide a "yes" or "no" answer. Consequently, the bioacoustician's question is not whether there is simply a statistically significant effect, which by itself may be meaningless and even misleading (e.g., Wasserstein et al. 2019), but whether the magnitude of the effect is biologically important. That is a much more difficult question to answer, and hence why it is often ignored albeit inadvertently. By ensuring that research questions have biological relevance, bioacousticians can design studies that can draw meaningful conclusions about animals and their populations.

Once the biologically relevant question has been identified, the bioacoustician can determine what study design is required and whether it is possible to carry it out. All too commonly, constraints occur in available budgets and time allocated to undertake the research. This often results in sub-optimal study designs and sample sizes (e.g., reduced numbers of surveys, available

C. Salgado Kent (✉)
Centre for Marine Science and Technology, Curtin University, Perth, WA, Australia

Oceans Blueprint, Perth, WA, Australia

Centre for Marine Ecosystems Research, School of Science, Edith Cowan University, Perth, WA, Australia
e-mail: c.salgadokent@ecu.edu.au

T. A. Marques
Centre for Research into Ecological and Environmental Modelling, University of St Andrews, St Andrews, Fife, UK

Departamento de Biologia Animal, Centro de Estatística e Aplicações, Faculdade de Ciências da Universidade de Lisboa, Lisbon, Portugal

D. Harris
Centre for Research into Ecological and Environmental Modelling, University of St Andrews, St Andrews, Fife, UK

© The Author(s) 2022
C. Erbe, J. A. Thomas (eds.), *Exploring Animal Behavior Through Sound: Volume 1*,
https://doi.org/10.1007/978-3-030-97540-1_9

acoustic instruments, and/or surveyed animals). The reality is that for a bioacoustician to be able to confidently answer research questions, budgets must allow for robust experimental designs and sufficient time to collect sample sizes representative of the study population. Even when budgets and time allow for carefully designed experiments, however, environmental conditions and study animals often cannot be controlled, particularly when studied in their natural environment. Moreover, many studies occur opportunistically and are not the result of an experimental design developed specifically for the study aims. They are observational in nature and can take advantage of large, long-term existing datasets or unexpected opportunities to collect field data. In fact, data collected opportunistically are prevalent in bioacoustical studies, as many researchers take recording systems into the field during other work to use when time permits.

The challenges described above, from ensuring that the research questions have biological relevance, to evaluating the achievability of a study and reliability of its outcomes, are only a few of many challenges faced by bioacousticians. To overcome these challenges, bioacousticians must have solid foundational knowledge about the quantitative aspects of their research: from how to formulate quantitative research questions, to designing robust studies and undertaking suitable analyses. Only by having these skills can reliable conclusions and scientific claims be made.

Today, not only are there a wide range of analytical tools available to select from, but this ever-increasing number has been evolving quickly over recent decades due to the dramatic improvement in computer capacity. Moreover, ongoing research in statistics continually updates our knowledge on the suitability of commonly used methods (Wilcox 2010). In some instances, methods previously used over a wide range of applications may now only be acceptably applied to certain scenarios, with new methods superseding old ones. Having said this, while a new method may be considered the 'Rolls Royce' of analyses, sometimes an older, simpler approach may still do the job well. Consequently, not only is it important for researchers to have a solid foundation in long-established analytical approaches, but they must keep up to date with new developments. In general, a researcher should understand the fundamentals involving randomness, variability, and statistical modeling discussed in this chapter, and be able to adapt them to their specific context—this understanding is arguably more valuable than a book of recipes that tells a researcher which method to use and when.

A consequence of the many advancements over recent years and the large range of analytical approaches available today is that selecting the right tool can be an overwhelming task. In fact, the right tool might not exist for a specific setting. In such cases, collaboration with an applied statistician may be fundamental. This chapter aims to give general guidance on considerations that bioacousticians should make when tasked with undertaking research resulting in what are often complex and messy bioacoustical datasets. The information presented in this chapter is by no means meant to provide a menu of analytical tools, their mathematical basis, or conditions of use. There are a large number of widely available textbooks that do just that, and many are referenced here. Bioacousticians should consult the relevant textbooks for in-depth knowledge of approaches, their applications, limitations, and assumptions about the characteristics of the data that must be met. Rather, the focus of this chapter is to provide practical guidance on: (1) the development of meaningful research questions, (2) data exploration and experimental design considerations (also see Chap. 3), and (3) common analytical approaches used today. The approach taken in this chapter is to define basic terms and concepts as they appear in the text, so that readers new to the subject can also understand the more complex concepts discussed, regardless of their prior statistical knowledge.

Note that this chapter has been written from the perspective of a biologist faced with the challenges common to bioacoustical research. If, from this chapter, the reader gains an appreciation of limitations in their data, considerations they

should make when selecting analytical approaches, and the biological relevance of their analytical outputs, then this chapter has achieved its purpose. Entire books could be written about how a bioacoustician, in fact, any ecologist, might become more quantitative. A good example of such a book is suitably named *How to be a quantitative ecologist* (Matthiopoulos 2010), which we wholeheartedly recommend as good reading after this chapter.

## 9.2 Developing a Clear Research Question

At the concept stage of any study, the purpose and specific research aim must be clearly defined. The research aim should be novel (i.e., not already answered in previous research). Once the general aim has been defined, the specific analytical research question can be developed. While developing the question may seem to be a simple, self-evident task, it requires careful consideration. The structure of the question drives the experimental design and selection of analytical tools, thus its accurate development is essential. To frame a question in clear, concise analytical terms, it is useful to identify the type of study involved. There are many types of studies conducted for a wide range of purposes. Depending upon the discipline, groupings that describe types of studies and their definitions vary. Here, we have adopted five of the six groupings referred to by Leek and Peng (2015) as common in bioacoustics. These study types include descriptive, exploratory, inferential, explanatory (called 'causal' in Leek and Peng 2015), and predictive studies. Definitions we give here have been framed within the context of common bioacoustical questions, and thus are adapted from more broad definitions.

Of the study types, descriptive studies are the simplest, aiming to summarize datasets collected. Exploratory studies take a step beyond and explore relationships, trends, and patterns in datasets. Neither of these types of studies attempts to infer beyond the dataset collected to the wider population. These types of studies are commonly used during preliminary data exploration before undertaking inferential, explanatory, or predictive studies (see Sect. 9.3.3). Indeed, descriptive and exploratory surveys are often used to develop the more complex inferential, explanatory, and predictive study type questions. Inferential studies build on descriptive and exploratory studies by quantifying whether findings are likely to be true for a broader population and hence can be generalized. For example, inferential studies are commonly used to make decisions about whether there is sufficient evidence regarding observed patterns or relationships in sample data to believe that they have not arisen from the population by pure chance alone. Explanatory studies aim to identify associated conditions (e.g., species, age, sex of an animal, date, time of day, season, and environmental factors such as temperature, noise, etc.) influencing or explaining an outcome (e.g., the rate at which animals produce their calls). These studies seek to determine the magnitude and direction of relationships (Leek and Peng 2015). Predictive studies aim to predict future outcomes in given conditions or scenarios (but may not necessarily explain conditions leading to an observed outcome). By identifying which of the study types your research aim falls into, the general structure of the analytical question can be formed. Some examples of the different study types and corresponding analytical questions are given in Table 9.1.

## 9.3 Designing the Study and Collecting Data

Once the analytical question has been formulated based on the study type, novelty, and whether it truly addresses the research question, the feasibility of collecting the required data will need to be assessed. Practical considerations, for instance, include identifying any hindrances to study site accessibility or timely ethics approvals and animal experimentation permits. Below (Fig. 9.1) is a checklist of some preliminary considerations before committing to developing, designing, and executing a study.

**Table 9.1** Examples of study types and their corresponding objectives and questions

| Study type | Purpose | Example objective | Example questions |
|---|---|---|---|
| Descriptive | Studies conducted to describe phenomena and conditions measured during a study. | Describe the characteristics of sound produced by sea turtle hatchlings recorded during a study. | • What is the frequency range of sounds produced?<br>• What are the source levels of sounds produced?<br>• What is the rate of sound production by sea turtle hatchlings? |
| Exploratory | Studies exploring relationships, trends, and patterns in datasets (not in a broader population). | Establish how observed hatchling sea turtles' sound production varied during a survey. | • How does observed hatchling sea turtles' sound production vary during a given survey? |
| Inferential | Studies aiming to estimate population parameters or test hypotheses about a broader population. | Determine the average expected sound production rate of a population of hatchling sea turtles. | • What is the average expected sound production rate of a population of hatchling sea turtles? |
| Explanatory | Studies that aim to understand the underlying cause(s) of a behavior, state, or phenomenon. | Identify what influences sound production in sea turtle hatchlings. | • Are communications influenced by the presence of other sea turtles, environmental conditions, or human/predator threats? |
| Predictive | Studies that aim to predict an outcome (such as animal behaviors) in response to a stimulus or condition. | Predict hatchling sea turtle sound production rate when threatened by humans. | • What will be the expected sound production rate of hatchling sea turtles when exposed to human threats? |

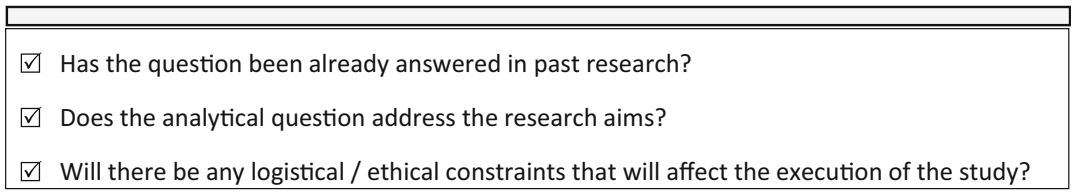



**Fig. 9.1** Checklist of some considerations to be made before committing to a study

### 9.3.1 Experimental Design

The ideal situation is to formulate the analytical question before data are collected (i.e., a priori) so that experiments can be designed to maximize the chance that, based on the observations, they produce precise (i.e., close to one another) and accurate (i.e., proximal to true values) estimates of the parameters of interest, and so that there is a high probability of detecting relevant effects (i.e., that there is sufficient statistical power) when they are present. In some cases, however, formulation of the analytical questions occurs after data have been collected (i.e., a posteriori). This may occur as a result of poor planning or of new and unforeseen research opportunities. A scenario in which this often occurs is when data already collected for another primary study are used to answer a new research question. In these cases, the methods and experiment are not necessarily designed according to the analytical requirements of the new research question. Bioacoustical studies using pre-existing opportunistic data often do so because collecting new data can be prohibitively expensive (e.g., if the field site is remote or if specialized equipment is required). Since the methods and experimental design may be sub-optimal for the current study questions, the data must be meticulously evaluated to check that newly formulated analytical questions can indeed be answered. Studies attempting to answer specific research questions using sub-optimal or poor-quality data cannot always be salvaged, even with sophisticated analyses. The prominent

twentieth century biostatistician, Sir Ronald Fisher, illustrated this problem with the following quote: "To call in the statistician after the experiment is done may be no more than asking him to perform a post-mortem examination: he may be able to say what the experiment died of" (Fisher 1959). This message cannot be overstated. It is critical, wherever possible, to consider the question carefully a priori, so that the study is able to answer the question (Cochran 1977). If you think you might need to consult with a statistician, do so before collecting the data.

For analyses to answer ecological research questions, the experimental design must yield sufficient information about the question of interest. Often, ecological questions involve sets of sampling units taken from a larger group (i.e., the statistical population, hereafter referred to as a population unless otherwise stated). For a given study species, or set of species, sampling units could be defined as individuals, groups, cohorts, communities, or local populations of the species of interest—it depends on the research question. Usually, due to logistical and time constraints, it is not possible nor desirable to make measurements over all objects or the whole population. In these cases, a sample is taken and data collected from the sample are considered to be representative of the population. It is key that the process used to draw the sample is well understood and is ideally random in design. The process of drawing conclusions regarding a population based on a sample from it is called statistical inference.

To make meaningful inferences about the properties of a population, the sampling protocol must yield a sample size that is sufficiently large to represent the population. In addition, the sampling protocol should either eliminate or control significant sources of error including random and systematic error (Cochran 1977; Panzeri et al. 2008). Random error is caused by unknown and unpredictable changes, such as in the environment, in instruments taking measurements, or as a result of the inability of an observer to take the exact same measurement in the same way. Statistical methods typically quantify this error and, in fact, build on it to draw inferences. In some sense, if there was no error then there would be no need for statistics. Of course, the performance of the analytical methods is affected by the amount of error in the data, in that the statistical power to detect significant effects decreases with increasing error, but if there was no error, by definition there would be no questions left to answer and statistics would have no role to play. Systematic error (i.e., bias) is consistent error that is repeatable if the data are recorded again. It can arise from many causes, such as a person consistently making the same erroneous observation (i.e., biased observation; e.g., incorrectly recording male birds as female birds) or an incorrectly calibrated instrument. In behavioral studies, biases in collected data can also be introduced by the presence of the researchers themselves (e.g., through human disturbance in a study on supposedly undisturbed animal vocal behavior). The introduction of bias can be further illustrated in the example of a bioacoustician estimating acoustic cue production rate (i.e., number of cues, such as calls, produced per unit time) for a population. In this example, the researcher obtains samples of animals by locating the animals producing acoustic cues. It is highly likely, however, that the sample collected will be only from animals that are in a sound-producing state (as silent animals will go undetected), hence acoustic cue rate might be inadvertently overestimated. Furthermore, animals may respond to the presence of the researcher by altering their cue production rates, thereby introducing further error to cue rate estimation. Such studies should be designed to remove or control biases. If controls cannot be integrated into the experimental design, then these may be able to be applied at the analytical stage (statistical controls; see Dytham 2011) and estimation of, and adjustments for, unavoidable biases may be made during the analysis. For topics on experimental design (e.g., systematic, stratified-random, and random-block) that aim to reduce biases and increase inferential power, the reader is referred to textbooks such as Lawson (2014), Manly and Alberto (2014), Cohen (2013), Underwood (1997), and Cochran (1977), among many others. It is critical that researchers carefully consider and identify the

☑ Does the scope of the experimental design match those of the questions?

☑ Is the sample size large enough given the effect size (see Section 9.5.1.2 for discussion on effect size) being investigated?

☑ Are the resources (e.g., time, money, and trained personnel) available for the project sufficient to carry out the study?

☑ Will data be reliable (i.e., accurate and precise) enough to answer the questions?

☑ Will causes of biases in data collected be able to be identified and removed or addressed adequately?

**Fig. 9.2** Checklist of some considerations to determine whether a research question can be answered

most suitable sampling design for their research questions.

Despite all attempts to obtain reasonable sample sizes, minimize biases, and carefully select an appropriate experimental design, data quality is frequently sub-optimal due to logistical or practical constraints. Often unexpected restrictive weather conditions and/or failure of instruments limit data collection during fieldwork. Good planning can mitigate unexpected data limitations, thus wherever possible, there should be contingency plans in place to deal with the unexpected (e.g., budgeting for a reasonable number of poor-weather days or redundancy in instrumentation). Even with careful design and contingencies implemented, data limitations can still occur and may need to be dealt with at the analysis stage. However, as noted before, sophisticated analyses to deal with these are always a second-best option over implementing data collection methods and survey design that are robust to potential limitations. Figure 9.2 gives a list of some considerations to be made for assessing whether research questions can be answered before data are collected.

### 9.3.2 Instruments and Measurements

Instruments must be able to measure subject behavior and conditions of interest in the study such that estimates derived from the observations have sufficient accuracy and precision to detect the effect(s) of interest. The accuracy of an estimate is its proximity to the true value, while precision refers to the variability of successive estimates of the same quantity. Naturally, to be able to derive accurate and precise estimates, measurements must also be accurate and precise. Accuracy and precision of measurements are evaluated through calibration and testing of the instruments. Some instruments may simply not have the capacity or range required for the study. For example, a low-frequency acoustic recorder will not have the capacity to measure the acoustic behavior of bats, which produce high-frequency echolocation signals. While careful consideration must be made in selecting instrumentation, considerable advances in their capacities have been made over recent decades. Instrumentation in bioacoustical studies is discussed in detail in Chap. 2. Below is a checklist for evaluating whether the selected instrumentation will collect the required data for a project (Fig. 9.3).

### 9.3.3 Preliminary Data Exploration

Data quality resulting from the experimental design, selected instrumentation, and measurements must be checked through data exploration and visualization (e.g., graphics, spectrograms) before embarking on planned

☑ Do the instruments have the sensitivity (i.e., sufficiently low noise floor and thus sufficiently low amplitude that can be recorded), dynamic range (i.e., range of amplitudes that can be recorded), frequency range (for sound recorders), and field robustness required for the study?

☑ Do the instruments obtain sufficiently accurate and precise measures?

☑ Is there a quality-control process to ensure that instrument accuracy and precision can be measured over time (e.g., systematic calibration and testing)?

☑ Are the instruments reliable in that they will not result in significant sets of missing or biased data?

**Fig. 9.3** Checklist of example considerations for selecting instrumentation for a bioacoustical study

analyses. It can be said that it is never early enough to explore data, nor can there be too many graphs involved in doing so. In fact, a preliminary exploration of data should always be conducted at the beginning of data collection to allow the structure of the data to be investigated, including the presence of anomalous data points, missing values, and potential biases. By identifying these early in the study, unforeseen design, sampling, or instrumentation issues can be rectified. Preliminary exploration of data, after data collection has been completed, will allow for any remaining anomalies and biases to be identified and planned analyses refined. Suspicious observations can be introduced at different stages of the research, for instance through: (1) data entry error, (2) changes in the measurement methods, (3) experimental error, or (4) some unexpected, but real variation. For the first three cases, the anomalous value(s) might be removed before analysis. In the last case, there could be some biologically important reason for the observed unexpected values. Sometimes the word "outlier" is used to refer to these suspicious observations, but we prefer to avoid the term. An outlier implies something that was unexpected, but only after defining what would be expected can we decide what the word "outlier" means. Often "outliers" are very informative and can even lead to new research questions. Consequently, it is important to understand how anomalies have occurred and to ascertain whether they should be removed or not. A good and honest approach, with little added cost, is to present and discuss the results of an analysis with and without those observations. This approach provides useful information about the practical consequences of the presence of anomalous observations.

If sufficiently large gaps in information from missing values occur, the data may not be representative of the larger population, especially since it might be hard to determine after the survey whether the data were missing at random. Similarly, if measurements were collected under certain conditions (e.g., poor weather or noise), the data cannot typically be used to make inferences outside this range of conditions (which would be referred to as extrapolation). Finally, data of very poor quality may not be salvageable, and—as mentioned before—it is far preferable to get the data right in the first place than to trust analytical solutions to deal with problems introduced at the data collection stage. Data exploration and visualization are further discussed in Sects. 9.4 and 9.5.

## 9.4 Data Types and Statistical Concepts

Regardless of the analytical approaches used, there are some fundamental terms and concepts that need to be understood before embarking on analyses.

### 9.4.1 Variable Types and Their Distributions

Measures of observations or conditions of interest in a study can be called variables. For instance, variables can be measurable properties of animals, their behaviors, or their environment. In a study of the acoustic characteristics of elephant vocalizations recorded at different ranges from the animal, relevant variables might include the range between the microphone and the elephant, the subject (i.e., which animal it is), the sound type, the received sound level, the spectral characteristics of the sound at the receiver locations, and the acoustic characteristics of the environment between the elephant and the receiver. In general, a researcher will have a good idea about the plausible values for the variables of interest, and hence what range of values to expect, but not know the exact values before the observations are made. Variables of known expected range but whose exact values are unknown until observed are random variables by definition. The notion of "outlier" is related to this expectation, as "unexpected" values might be considered suspicious. Within a regression context (see Sect. 9.4.3 for more detail), the variables that represent the outcome of interest are called dependent variables or response variables. When they represent the conditions that influence the outcome, they are called independent variables or explanatory variables, sometimes known as predictors or covariates. Hereafter we use all terms to discuss variables, choosing each time the definition we feel will help to make the meaning of a concept most intuitive.

Variables can be of two types: (1) categorical, which can be further subdivided into nominal or ordinal (if there is an order), and (2) numerical, which could be discrete or continuous. Categorical variables are often called factors and are qualitative. For example, if the variable was a sound type produced by a bird categorized as either song or chirp, then sound type would be a nominal factor with two levels, also called a binary variable. If the bird species was known to produce three different sound types, then the corresponding factor would have three levels. Numerical variables are quantitative, and can be discrete (e.g., integers such as counts) or continuous (where, by definition, an infinite number of values are possible between any two values). Examples of continuous variables are the height and weight of an individual or pressure and temperature, while the number of sounds or the number of individuals are examples of discrete variables. A summary of variable classification and metrics is given in Table 9.2.

Properties of these variables, such as central tendency measures like the mean, mode, and median, or measures of spread like variance and standard deviation, are statistics that can be used to describe a sample of values. When these refer to the values that these quantities have in the population (as distinct from a sample of that population), these properties are called parameters.

Often, additional variables are collected that are not necessarily of interest in explaining a research question but could influence the response variables. For example, while a bioacoustician might be interested in measuring the rate of vocalization of chicks as a function of the parents' presence, the frequency of predator visitation could also influence vocalization rates. In this example, collecting information on the main independent variable (parent presence) and the variable not of direct interest (predator presence) would be considered important to capture all variables influencing vocalization rate. Some of these variables might be of direct interest, but some might just be included in a study because they can affect the response, and if ignored, would confound the results. For this reason, they might sometimes be referred to as confounding factors or confounding effects. Note that these terms and their definitions vary with discipline (e.g., there is some discussion about the exact definition of a covariate; see Salkind 2010) and analytical software, and sometimes are used interchangeably. Therefore, the reader should make sure that, when reading a source or when reporting their own results, the context provides the required clarity for the wording chosen.

Not only are variables described according to the properties they measure and whether they are

**Table 9.2** Variable classification and metrics

| | Categorical | | Numerical | |
|---|---|---|---|---|
| | Nominal | Ordinal | Discrete | Continuous |
| Description | Non-ordered categories | Ordered categories | Variables in which the data can take on only certain values (i.e., values that have non-infinitesimal gaps between them containing no values) | Variables in which the data can take on real values and all infinitesimal real values between them |
| Example | Sound type (e.g., downsweep, upsweep, constant tone) | Vocal activity on a scale ranging from not vocally active (0) to highly vocally active (5) | Acoustic cue counts | Received sound exposure level (in dB) |

independent or dependent variables, but in the context of some analytical methods (e.g., linear regression models and their extensions) they are also described by whether they represent a specific or random set of values. Generally, in statistics, a variable with a value that is not known before it is observed (e.g., peak frequency of a call or number of animals in a group), but of which the range of possible values is known (e.g., a positive continuous number like the amplitude of a lion's roar), is known as a random variable, as described above. Its range of possible values is referred to as the domain of the random variable.

A random variable can be characterized by its probability distribution, which describes the probability of observing values in a given range of the domain of the variable. An infinite number of distributions exist, but some, given their useful properties, are widely used. These distributions are given names so that we can easily refer to them. Arguably, the most widely used are the Gaussian distribution (perhaps more often known as the normal distribution, but since there is nothing normal about it and it induces practitioners to think there might be, we avoid the term here), gamma distribution, and beta distribution, used to model continuous data; while the Poisson distribution, negative binomial distribution, and binomial distribution are useful when modeling discrete values. The uniform distribution is one in which all values in the domain are equally likely and can be either continuous or discrete. These distributions are typically defined by their parameters. As an example, the normal distribution is defined by the mean and the standard deviation, and for the case of the Poisson, it is defined by the mean only. Given the parameter values that define a random variable, all the characteristics of the random variable are unambiguously defined.

Values of a discrete variable are characterized by a probability mass function (pmf). A pmf is a function that gives the probability that a single realization of the variable takes on a specific discrete value. The number of vocalizing individuals detected in an area might be approximated by a Poisson random variable, characterized by its mean (such as 3.7 individuals). The Poisson distribution is special in that its variance is equal to its mean, a restriction that means that often it does not fit biological data well, where larger variance than the mean is the norm.

In contrast, continuous variables can be characterized by a probability density function (pdf). In the instance of a variable such as the change in duration of song, the pdf might be represented by a Gaussian distribution—a bell-shaped curve characterized by its mean and standard deviation. For example, the variable "change in song duration" could have a true mean change in duration of 240 s and a true standard deviation of 12 s. These true values are generally unobserved, but we would like to estimate them. A single measurement of change in song duration by a researcher could produce a value of 228 or 271 s. These single values are referred to as realizations of the random variable. Pdf functions provide information about how the values are distributed before they are observed. Further

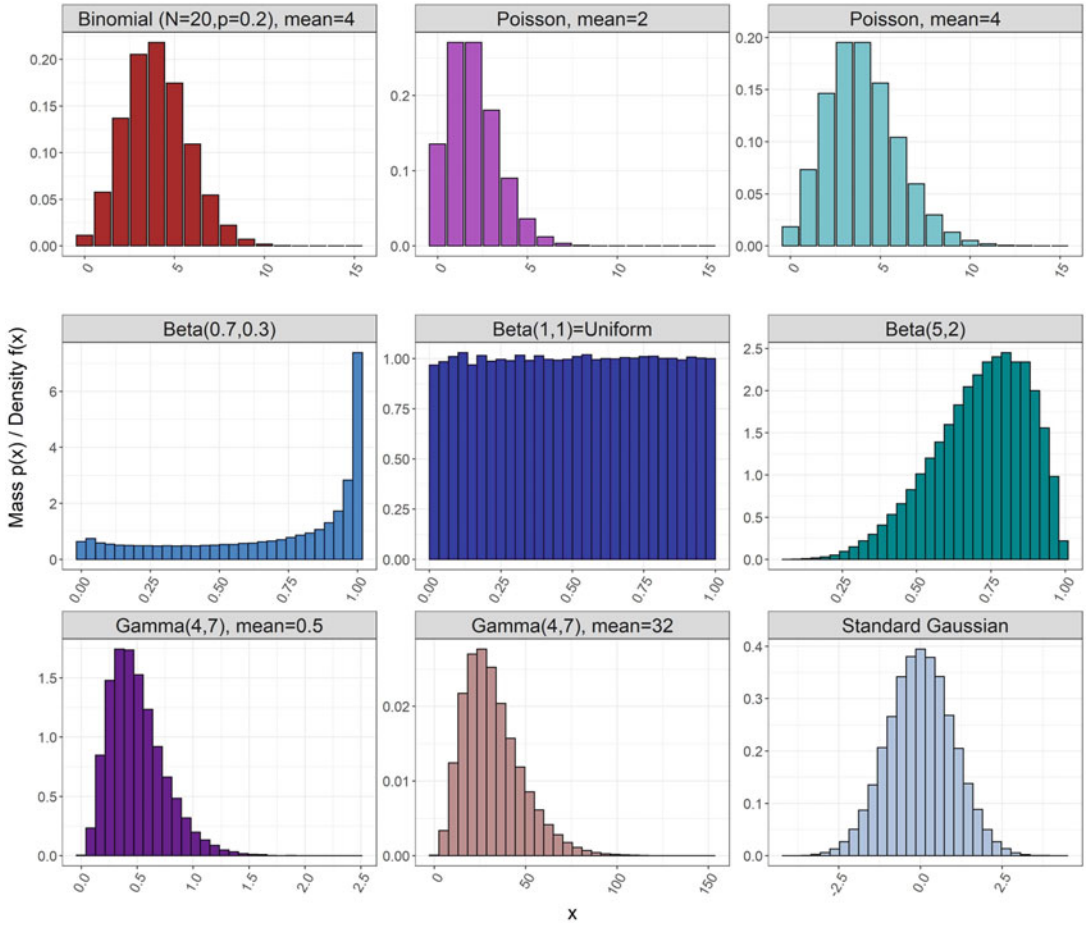



**Fig. 9.4** Examples of samples taken from different distributions. The Gaussian, gamma (defined by its shape parameter $k$ and scale parameter $\theta$) and beta (defined by shape parameters $\alpha$ and $\beta$) are continuous distributions, represented with histograms. The Poisson (defined by its mean) and binomial (defined by $n$ independent experiments and outcome success probability $p$), represented with barplots, are discrete distributions. Note some distributions can be special cases of others. As an example, the beta distribution, with shape parameters $\alpha = 1$, $\beta = 1$ is shown, illustrating the fact that it is equivalent to a uniform distribution

examples of distributions are given in Fig. 9.4. The reader is referred to Quinn and Keough (2002) for a good introduction to useful probability distributions in biostatistics.

### 9.4.2 Estimators and Their Variance

In this section, we introduce estimators and related concepts because we will need them later, but we note that we do so very briefly, just so that the terms do not come as a surprise. The reader is referred to Casella and Berger (2002) for further details on statistical inference, estimators and their variance.

As discussed previously, a parameter is a quantity relating to the population of interest. When performing statistical inference, we want to estimate the parameters in the population (e.g., the mean cue production for a species of whale) using samples (e.g., a sample of acoustic tags put on whales). To estimate parameters, we use

estimators. An estimator is a formula that we can use to compute a parameter based on a sample. In the case of estimating the population mean, the estimator is, not surprisingly, the well-known formula for the sample mean. Estimators are therefore based on random variables, in the sense that each time we collect a sample we would get a new observed value (i.e., a new estimate). Thus, an estimator can also be thought of as a sample statisitic that estimates the population parameter such as the mean. If we collected infinite samples and computed the estimator each time, we would get the estimator sampling distribution, from which we could evaluate the bias and the variance of an estimator. However, collecting infinite samples is not possible, but by understanding the properties of the estimator and the design used to collect the data, we can also quantify the variability associated with an estimator, based on a single sample. Variability is a key attribute of an estimator, and the resulting estimate from the single sample (known as the point estimate) is not enough to provide a full representation of it. For example, it is very different to say that we estimate a cue production rate to be 7.2 sounds per hour, than to provide the additional information that it could vary from 7.1 to 7.2, or that it could vary from 1.2 and 27.7. In the first example we have a small variance, and the latter we have such a large variance that the estimator itself is borderline useless. To compute an estimator's variance, there are two main approaches. If the estimator and the process by which we collect the sample is simple enough, we have standard formulae for the variance. That is the case for the sample mean from a simple random sample. However, often in practice, that is not the case, say because the sampling procedure is convoluted, there is a hierarchy in the process, or the estimator is composed of several random components, possibly not independent among themselves. A good example is an animal density estimator from Passive Acoustic Monitoring (PAM), where different random components like encounter rate, detection probability, cue rate, and false-positives might be at play (see Sect. 9.6.2 for a PAM density estimation example). In such cases, resampling techniques like the bootstrap might be considered. The rationale behind the bootstrap is that one can resample with replacement from the original sample, and the variability of the estimates computed over the resamples is an estimate of the estimator variability. The reader is referred to Manly (2007) for further details about these procedures. While variance is commonly reported, when comparing variances of quantities that have different means, the coefficient of variation (CV), which is the standard deviation divided by the mean, can be useful. The CV is typically reported as a percentage (% CV = standard deviation/mean $\times 100$).

### 9.4.3 Modeling

In its most simplistic form, a model is a mathematical generalization of the relationship among processes (Ford 2000). Models are by necessity a simplification of reality. Extending a quote popularized by George P. Box (1976), all models are strictly wrong, in that they are always oversimplifications of reality, but many models are useful, in that they provide useful explanations or predictions of reality. Models can either be empirical or theoretic. A common example of a theoretical model in acoustics is the piston model used to represent the beam pattern in a directional sound source like the dolphin biosonar system (Zimmer et al. 2005). While theoretical models are based on theory, empirical models are based on observations. Here we will focus discussion on empirical models as observed data are commonly used to fit models to describe bioacoustical processes. Models describing the relationships between whale vocalization rates and season or location (Warren et al. 2017) or dolphin occupancy and pile driving noise (Paiva et al. 2015) are examples of empirical models. Another example is a mathematical equation that describes the number of bird calls recorded within a given period as a function of the number of birds present. By identifying the mathematical relationship between variables, past events can be explained and future scenarios predicted. However, finding such an association requires careful interpretation, especially in observational

studies. Finding an association between two or a set of variables does not necessarily imply a causation. This could be either a spurious association, or an observation induced by a variable that was not recorded. It is a statistical capital sin to confuse correlation with causation. For example, on hot days, the consumption of ice creams increases, and so does the number of fires. But you can eat an ice cream guilt-free as you will not cause a fire!

#### 9.4.3.1 Introduction to Regression: The Cornerstone of Statistical Ecology

Arguably, the most common and most useful class of statistical models are regression models. The simplest regression model (i.e., the Gaussian linear regression model) has three basic components: (1) a dependent variable that is to be modeled (i.e., described or explained), and (2) independent variables that are thought to influence the dependent variable. The third component, the random error, distinguishes statistical models from deterministic mathematical models. The random error captures how the model differs from the actual observations. In other words, it measures how well, or badly, our model describes reality. Written as a mathematical expression, the simple regression model looks like this:

$$Y = \alpha + X\beta + \varepsilon, \tag{9.1}$$

where $Y$ is the response variable, $\alpha$ is the intercept (a constant), $X$ is the fixed independent variable, $\beta$ is the regression coefficient for the fixed independent variable that describes the rate of change of the response variable as a function of the independent variable, and $\varepsilon$ is the random error. In general, the parameters $\alpha$ and $\beta$ are not known and must be estimated based on data.

Most variables, particularly in ecology, are influenced by many covariates, and hence models can include multiple independent variables. For instance, in a study on whether the vocalization rate of sea lions differs with sex and age, vocalization rate (i.e., number of vocalizations per unit time) would be the response (dependent) variable and sex and age the explanatory (independent) variables. In addition to having these two explanatory variables of direct interest, other variables may also be relevant to include in models, because they might a priori be expected to also influence the response variable. Variables that may affect vocalization rate may include time, season, social context, or location. Studies in which multiple explanatory variables influence the outcome might have interactions between the explanatory variables that are important to consider. For instance, vocalization rate may differ between male and female sea lions, but only for sub-adults and adults and not for pups and juveniles.

In a regression model, a distribution is typically assumed for the response variable. This will induce a distribution for the random errors. Historically, regression models considered the errors of the dependent variable to be Gaussian distributed, and much of regression theory was developed under this assumption. Note that a model assuming a Gaussian error distribution in the dependent variable is commonly simply referred to as a linear model. Nowadays many generalizations to linear models exist (as described below and see Zuur et al. 2009 for common examples in ecology; see Generalized Linear Models in Sect. 9.5.3 below). Arguably, as noted above for random variables, the more commonly used distributions in regression models are Gaussian and gamma for continuous data, Poisson and negative binomial for counts, binomial for binary data, and beta for proportions (or probabilities), but many others exist. As for linear models, generalizations assuming other distributions associated with the response variable and associated error structure are commonly referred to by their distributions. For example, a Poisson distributed response variable with associated error structure of counts of animals is commonly referred to simply as a Poisson model. A gamma model might be used to model continuous positive values resulting from measurements of duration of a recorded song. Values representing the probability of producing a sound (between 0 and 1), however, might be modeled assuming a beta distribution.

Regardless of the error distribution of a model, classical regression models assume that observations are independent of each other (i.e., the value that one observation takes on is not influenced by another). The easiest way to ensure this happens is by design, and all efforts should be made to enforce it. In the biological world, the assumption is very often violated, and almost as often ignored. This can lead to errors in inferences made, the severity of which depends upon the degree and type of non-independence between observations. A few obvious sources of lack of independence (i.e., dependency) are observations collected within groups that share a characteristic (e.g., a litter or a pod of animals), or observations collected over space (where two observations are more likely to be similar the closer they are in space) and over time (where two successive observations are more likely to be less independent than two observations separated by a longer period of time). Researchers often mistakenly analyze data collected without proper consideration of whether observations are independent. By exploring and accounting for dependencies, or even purposefully including them in an experimental design, the power of an analysis may be enhanced. As an example, in a repeated measures study of bird vocalization rate as a function of time of day, repeated measurements of the same individuals during the day and night could be undertaken by design (instead of randomly sampling birds at each time period). Another example is that of a chorusing group of insects, in which sounds can be produced for hours. A researcher may be interested in measuring whether the insects chorus in a given 5-min period. At any point of time within a chorusing bout, the probability that insects will be chorusing in a 5-min time window will be expected to be high if they were chorusing during the previous 5 min. This leads to what are called autocorrelated observations. In such cases, the autocorrelation structure can be incorporated into the model. If evaluating the effect of time was not of specific interest in this study, an alternative and simpler solution would be for the model to use subsampled data to include only times at which insect sound production can be considered independent. However, by explicitly accounting for the autocorrelation structure in the model, more efficient inferences are bound to be obtained as there is no loss of information. Model implementation does become a bit more complex, however. Studies that purposefully measure subjects or populations repeatedly over time to create a time series of data are called longitudinal studies. Because time-series measurements, such as those from longitudinal studies, usually cannot be considered independent from one another (e.g., an animal's current behavior is likely dependent on its behavior during the previous sample time), a wide range of models have been purposefully developed to account for non-independence (see Sect. 9.5.3). Researchers should carefully consider and plan for potential sources of dependency in the design of their studies and data collection protocols.

A checklist of some considerations for describing and defining variables in your study, including whether they are autocorrelated or not, is illustrated in Fig. 9.5. These considerations should be made as part of the experimental design and analytical planning process prior to data collection and will need to be reassessed post data collection.

## 9.5 Tackling Analyses

In this section, common analytical approaches used in descriptive and exploratory studies are presented first, followed by those used in inferential, explanatory, and predictive studies. It is important to note that analyses relevant to inferential, explanatory, and predictive questions require preliminary data exploration (see Sect. 9.3.3), thus requiring descriptive and exploratory analyses first. In these cases, preliminary exploration of data attributes may refine previously planned analytical approaches. This is particularly relevant since sufficient data quality and specific distributions are required for empirical model assumptions to be met and these features can be assessed via initial data exploration.

Analytical approaches described in this section are examples only of a wider range available. The

☑ Have the variables and variable types been identified?
☑ If there are multiple independent variables, are there interactions of interest and/or are any of variables highly correlated?
☑ Are data for variables likely to be independent or autocorrelated?

**Fig. 9.5** Checklist of some considerations for defining variables in your study

purpose is, by way of examples, to provide a taste of the explosion of tools developed over the past few decades, the lively discussion that has arisen from their varied and inherent limitations, and the resulting developments in statistical approaches. The reader is directed to the wide range of available statistical textbooks and scientific papers to gain an in-depth understanding of the full range of approaches, their underlying concepts, and their correct use, limitations, and interpretation of outputs.

### 9.5.1 Descriptive and Exploratory Research Questions

Having defined the question (Sect. 9.2) and identified the variable types and some of their attributes (Sect. 9.4), tackling the analyses is the natural next step. For descriptive and exploratory questions and preliminary data exploration, summary statistics and graphical visualizations provide information about the attributes of variable measures and patterns and relationships in data. The information relates only to the properties of the observed data. Analyses that aim to generalize a sample to a population require inferential, explanatory, and predictive type analyses (discussed in Sects. 9.5.2 and 9.5.3).

#### 9.5.1.1 Univariate Summary Statistics and Graphical Visualization

Exploration and visualization in their simplest forms are undertaken by evaluating each variable on its own (Fig. 9.6). Analyses of single variables are called univariate analyses and are used for representing and summarizing the characteristics of the variable in question. For example, univariate exploratory statistics describe a variable's properties such as statistics for central tendency including the mean (note that there are different types of means; e.g., arithmetic, geometric, and harmonic), median, or mode, and spread of data including the range (maximum and minimum), variance, standard deviation, skewness (degree of asymmetry), kurtosis (i.e., how peaked a distribution is), or interquartile range (see Table 9.3). Data corresponding to a single variable can be summarized and explored using a range of graphing tools, such as histograms, box plots, bar charts, or scatterplots. Additionally, geographical data can be explored on maps and marine charts, and acoustic spectral characteristics on spectrograms (representing signal strength over different frequencies over time). As noted previously, it is (arguably) almost impossible to produce too many graphs at an exploratory stage—the more that you can learn about your data, the better. The reader is referred to standard statistical textbooks for information on the large range of summary statistics and graphical visualizations available (e.g., Zuur et al. 2007; Zuur 2015; Rahlf 2019 for examples in R).

#### 9.5.1.2 Bivariate and Multivariate Descriptive Statistics

The analyses of two variables together are called bivariate analyses. For instance, exploration and visualization of a given variable as a function of another variable to investigate possible correlation is a bivariate analysis (see Fig. 9.7). A practical example of a bivariate visualization is the use of box plots to visualize the distribution of call types (one variable) as a function of age class (a second variable), or a scatterplot of a recorded acoustic cue rate as a function of time of day. Following this logic, multivariate analyses

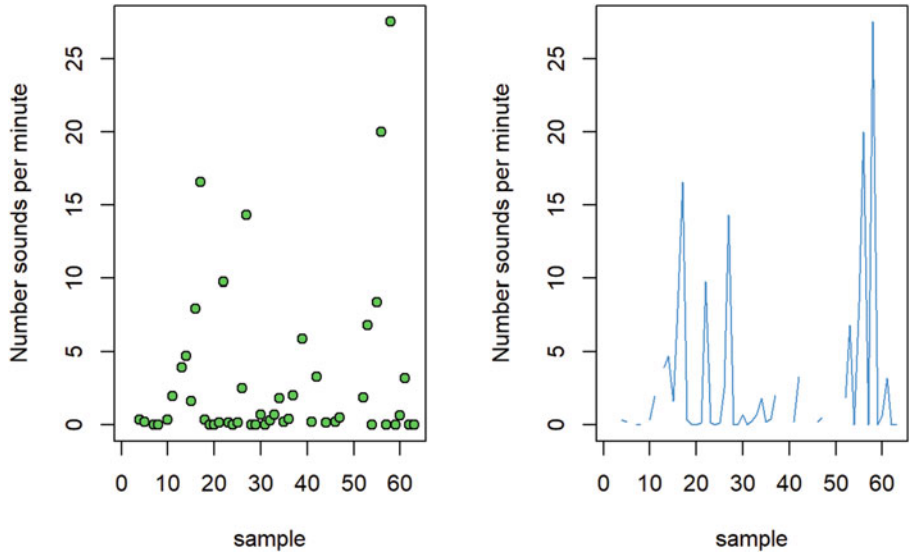



**Fig. 9.6** Example of univariate data visualizations of dolphin sounds detected: (left) scatterplot and (right) line chart. Data source: WAMSI as part of Project 1.2.4 (Brown et al. 2017)

**Table 9.3** Description of example univariate analytical and visualization tools

| Measure | Statistic | Visualization tools | Common purposes |
|---|---|---|---|
| Location and central tendency | Mean (arithmetic, geometric, harmonic), median, mode | Point, line and bar charts, histogram, boxplot | Describe the central tendency of values in a variable |
| Spread | Range (maximum and minimum), variance, standard deviation, skewness, kurtosis, standard error, interquartile range | Scatter plots, box plots, interquartile range, point, line and bar charts with standard error bars | Describe the spread of measures in a variable and identify patterns and data gaps |

naturally consist of the joint analysis of multiple variables. Visualization tools and summary statistics can also be applied to multivariate analyses. For instance, two and three-dimensional scatterplots, bar charts, stacked bar charts, and multiple line graphs can display statistics and spread of data as a function of multiple variables on the same figure.

When bi- or multivariate analyses aim to explore associations and patterns, the magnitude of the association can sometimes be quantified. For example, in a bivariate analysis, the magnitude of the linear relationship between two variables can be quantified using a statistic called Pearson's correlation coefficient ($r$). The magnitude of an association such as this one is often referred to as an effect size. For example, Pearson's correlation coefficient is a standardized metric ranging from $-1$ to 1; with a perfect negative association yielding a value of $-1$, no association 0, and a perfect positive association a value of 1. In some disciplines, conventional criteria have been suggested to classify effects as small, medium, and large (see Cohen 1988). What may be in one study considered a large effect (say, $r = >0.6$), however, may not necessarily be in another study (where say, $r = >0.8$ might be considered large). Consequently, evaluating what is a meaningful effect size that a study aims to detect should always guide the design of a study and interpretation of its outcomes. It is a question that the researcher

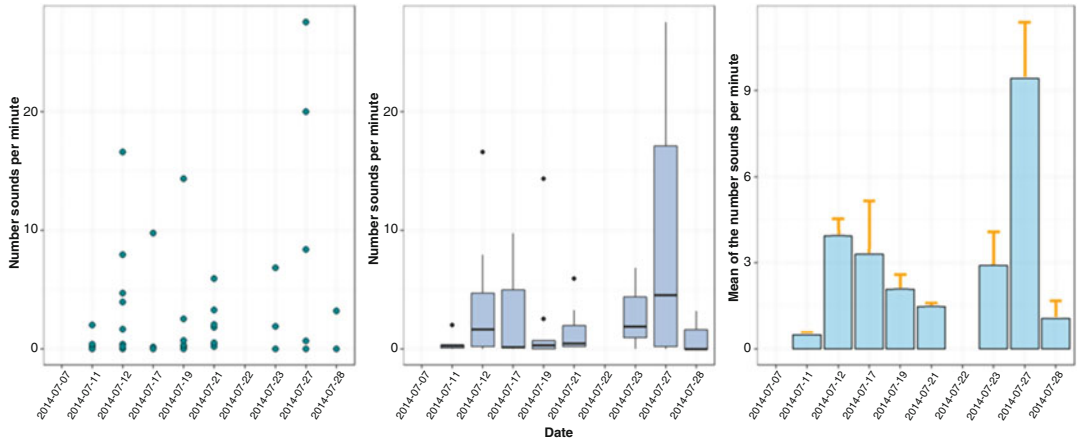



**Fig. 9.7** Example of bivariate data visualizations of dolphin sounds detected during July 2014: (left) scatterplot, (middle) box plot, and (right) bar chart with standard error bars. Data source: WAMSI as part of Project 1.2.4 (Brown et al. 2017)

should answer based on their biological knowledge and is not related to statistical considerations.

When a study's goal is to explore associations and patterns among many variables, analyses become more complex. Multivariate approaches are commonly used to reduce many variables to a few key ones. This is known as dimension reduction. Multivariate approaches are also used to explore relationships and clustering, and to classify objects based on common multiple variable attributes. A good source for additional details on multivariate methods is Borcard et al. (2011).

One of the most common analyses used for dimension reduction is principal components analysis (PCA). The name of the method is derived from the fact that new variables, known as principal components, are obtained from the set of original variables. For example, a researcher may be interested in exploring whether populations of a social insect, such as a species of ant, can be determined based solely on acoustic signals (e.g., stridulations) its individuals produce for communication. In this case, a range of variables might be measured, such as pulse duration, bandwidth, minimum and maximum frequency, and intensity, to name a few. In acoustics, a large number of variables might be measured to capture the full range of characteristics of acoustic signals. Consequently, using a data reduction method to capture the most variance explained by these variables by creating just one or two new variables (called principal components in PCA) makes the exploration of patterns in sound characteristics easier. The first principal component retains most of the original variance, followed by the second component, and so forth. These principal components are sometimes called factors. Factor 1 and 2 can be plotted against each other, and distinct groupings of plotted values for different populations would be suggestive of differing characteristics in stridulations among populations. To statistically test differences, PCA might be used to generate factor scores as inputs into inferential, explanatory, and predictive analyses (e.g., a regression analysis). Note that there are many dimensionality reduction approaches (see Van der Maaten et al. 2007), and researchers planning on using these tools should acquaint themselves with the wide range available today, their conditions of use, and their limitations. While one approach may be suitable given the attributes of one dataset, another may be required for a different dataset.

Clustering and classification analyses assign objects into groups based on measured attributes (variables). Cluster analyses form groups (McGarigal et al. 2000; Zuur et al. 2009) using

"unsupervised learning," where you do not "train" the procedure by labeling "training" data with group membership as you might in other methods. A range of cluster analysis algorithms are available including common approaches such as k-means and hierarchical clustering (see Borcard et al. 2011). Clustering and classification are used commonly for pattern recognition and are described further in Chap. 8.

Many other multivariate analytical approaches are available, ranging in their assumptions, strengths, and limitations, and the variable attributes for which they are most suitable. For example, correspondence analysis (CA) is similar to PCA, but can better cope with categorical data. The reader is referred to the many textbooks on the subject, such as Everitt and Hothorn (2011) on some of the more commonly used multivariate methods and their practical application in the software R.

As in the univariate case, we reiterate that associations identified in exploratory multivariate analyses do not indicate causation. Researchers interpreting exploratory analysis results should take care to never conclude that the results are evidence of causation. A brief checklist has been provided below with examples of the types of data considerations required for selecting analyses suitable for descriptive or exploratory questions (Fig. 9.8). The checklist is not exhaustive, rather it is indicative of the kinds of considerations required.

### 9.5.2 Inferential Studies

Statistical inference is used to infer properties of a population (e.g., estimate parameters) or test hypotheses. There are two widely used distinct frameworks for making statistical inferences: the frequentist and the Bayesian paradigms. Classical frequentist inference has a long history and has dominated past animal behavior and ecology research, while Bayesian inference is becoming increasingly popular. Both approaches can provide insightful information, however, they represent different interpretations of probability.

In frequentist probability, the probability of an outcome occurring is based on the relative frequency of occurrence based on a large number of observations taken. For example, the probability of bird vocalizations being recorded at a study site might be based on many sample recordings taken under the same conditions at the site. If vocalizations occurred 48% of the time, the probability of the outcome of birds vocalizing would be interpreted as 0.48. As the sample size increases, the proportion of occurrences approaches the true (unknown) proportion. If the sample size is small, the calculated proportion may not be a reliable representation of the true probability.

In the Bayesian interpretation, the probability is the degree of *belief of the likelihood* of the outcome. For example, it may be that a researcher believed that vocalization in nesting birds is related to predator presence. The researcher had visited the site and rarely heard birds vocalizing when predators were absent but noticed them vocalizing more often when predators were present. Maybe the researcher had even made a few recordings when predators were present and absent and found that birds were vocalizing 5 out of the 10 times she recorded in the presence of predators and 1 out of 10 times in their absence. In this example, these observations would constitute the prior belief. The research then undertakes a study designed for the purpose of collecting an unbiased set of observations to be used in analyses (sampling in the presence and absence of predators). Using Bayes' Theorem, the prior knowledge can be used to calculate the probability of vocalization that accounts for knowledge before and after collecting evidence (sampling). If the number of samples is large, the resulting probability estimate may not change much from that obtained in a frequentist framework. However, if the sample size is small, the prior knowledge may significantly affect the estimate of probability. Therefore, the lower the sample size (i.e., in general the lower amount of data coming from the data), the more the prior becomes important.

Many professional statisticians fall firmly in the frequentist or Bayesian camp. This often

☑ Do I require description, exploration, and visualization of individual variables, either for answering the main study question or for checking the quality of the data and assumptions of analyses planned for inferential, explanatory, or predictive studies?

The answer to this is always **YES**. Data always need to be checked for quality and attributes, and if the question requires inference or empirical models, the validity of assumptions needs to be checked (see Section 9.3.3 and 9.4)!

☑ What types of variables do I have?
☑ Does the study question involve single or multiple variables?
If multiple variables,

☑ are there a large number of variables that I need to reduce, explore their association, or investigate clustering or classification of groups characterised by them?

**Fig. 9.8** Checklist of some considerations for identifying approaches for descriptive and exploratory questions

follows directly from their training, or just by convenience and actually not having thought much about the philosophical ramifications of their choice. Sometimes they are rather inflexible in their beliefs (be it in one or the other camp). We recommend a more pragmatic approach in practice. Depending upon the problem at hand, one or the other framework might be more suited to the question, easier to implement, or more sensible for incorporating all available information (Nuzzo 2014; Ortega and Navarrete 2017). Consequently, we believe that the modern bioacoustician should have a basic understanding of the differences between frequentist and Bayesian approaches, and suggest that rather than only being frequentist or Bayesian, a pragmatic approach be taken. Below, we provide a very brief introduction to statistical inference applied to parameter estimation and hypothesis testing.

#### 9.5.2.1 Parameter Estimation

There are a range of approaches to estimate population parameters, such as the population mean or variance, or a shape or scale parameter of a distribution, from a sample. In the context of ecological modeling, the frequentist approach to estimating parameters typically uses maximum-likelihood (Hilborn and Mangel 1997). In Maximum Likelihood Estimation (MLE), parameter values of a distribution are estimated by maximizing the likelihood function so that the MLE estimates are the values of the parameters that are most likely given the sample data. An alternative method is Least-Squares Estimation (LSE), where a solution that minimizes the sum of the squares of the residuals (the difference between the observed values and those obtained using the fitted model) is obtained. For a Gaussian-distributed response variable, and several other simple examples, the LSE solution is equivalent to the MLE. Nowadays LSE are mostly introduced for teaching purposes, and most implementations use maximum likelihood.

As indicated above, the Bayesian framework combines information on the likelihood of an outcome using observed data with prior information on the distribution of the unknown parameter being estimated. The prior distribution can be an assumption based on the researcher's understanding and experience of the parameter before the study began or it can be based on the results from a pilot or previous study. Often the prior distribution simply reflects a lack of knowledge and may be uniform over all the possible values the parameter of interest might take (i.e., the parameter space). A posterior distribution (i.e., updated understanding) is attained by multiplying the prior distribution function with the likelihood

function and scaling the result to provide a probability distribution function. All the inferences are then based on this posterior distribution. The posterior distribution thus can be seen as a compromise between the prior information and the information contained in the data, expressed via the likelihood function. There are various resources available for further reading on the Bayesian framework. Ellison (2004) provides an excellent and gentle introduction to the use of Bayesian methods in ecology, while McCarthy (2007) provides a more thorough overview. Stauffer (2007) gives an in-depth introduction to Bayesian and frequentist statistical research methods and Gelman et al. (2013) discuss Bayesian data analysis. *Statistical Rethinking* by McElreath (2020) is a comprehensive treatment for a reader wanting to become fully versed in the Bayesian philosophy, including R code to explore all the key concepts.

When inferential methods, such as those introduced above, are used to estimate parameters from sample data, the inferences we draw from them are uncertain. Confidence intervals (CIs; a frequentist approach) and credible intervals (CrIs; Bayesian counterparts) are tools for expressing our uncertainty about parameter estimates. Confidence intervals, although more widely used, are arguably more difficult to interpret than credible intervals. Confidence intervals give information based on our sample estimate, and by definition, if we repeated the procedure many times, 95% would include the true parameter value. Note a 95% CI does not mean that 95% of the observations lie within the interval, nor that the probability of the true value of the parameter being in the estimated interval is 0.95. After you estimate the confidence interval, the true parameter value either is, or is not, in the interval, even if we do not know which it is. In contrast, 95% CrIs would represent a range of values for which there is a 0.95 probability that the parameter falls in that range. Ironically, what this means is that while most people use frequentist confidence intervals, they often interpret them, incorrectly, as credible intervals. Although credible intervals are intuitively easier to understand, they can be more difficult to calculate than confidence intervals.

#### 9.5.2.2 Hypothesis Testing

While hypothesis testing has been traditionally undertaken using a frequentist approach (called null hypothesis significance testing, NHST), equivalent Bayesian approaches are increasingly applied. This section focuses on providing a brief introduction to NHST as a foundation and provides references for further reading on Bayesian approaches. These basic concepts are introduced here with examples of their application to test statistics (i.e., statistics values used to reject or support a null hypothesis), however, they are also an integral part of modeling and model selection in explanatory and predictive questions (discussed in Sect. 9.5.3).

NHST constitutes a widespread paradigm under which research has been conducted (NHST, Fisher 1959), however, it is often not used sensibly, and frequently blindly used and abused. In some of these cases, pressure on researchers to find statistically significant effects has resulted in poor research practices (see Nuzzo 2014; Beninger et al. 2012 for detailed discussions on the topic). Applying NHST to reasonable hypotheses and qualifying results according to the limitations and assumptions of NHST, however, can produce important new knowledge. To achieve this, an understanding of how NHST works is required. Here we provide insight into the framework by way of example.

Under the NHST framework, researchers put forward a hypothesis (i.e., proposed explanation) about the phenomena being studied based on a study question. Let us say the researchers' question is "Do seal pup call rates differ between night and day?" The null hypothesis ($H_0$) is that call rates do not differ between night and day, and the corresponding alternative hypothesis ($H_A$) is that pup call rates do differ between night and day. Note that this hypothesis implies a two-tailed test, one for which the null hypothesis is rejected if a positive or a negative effect (i.e., a large or small value of the test statistic) is found. In contrast, a one-tailed test would be used by a researcher interested only in the difference between groups in a specific direction (e.g., "Are call rates greater during the day than at night?").

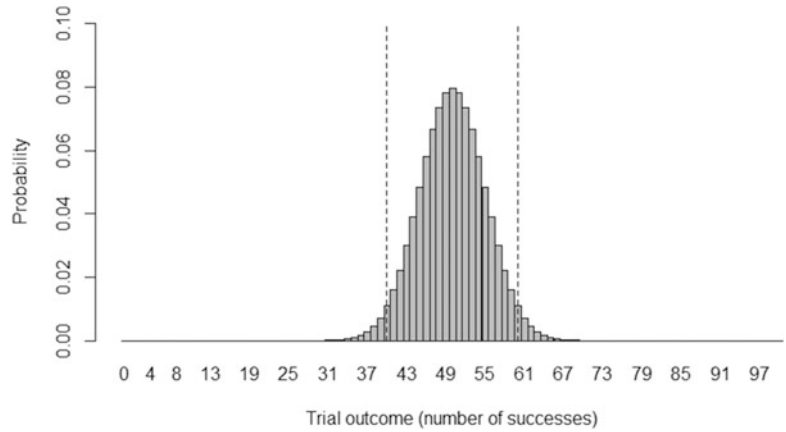



**Fig. 9.9** Binomial probability mass function with parameters $n = 100$ trials and $p = 0.5$, with the quantiles 2.5% and 97.5% represented by vertical dashed lines. Under $H_0$ only 5% of the observations would be more extreme than those quantile values

In this example, the researchers cannot measure the call rates of all animals in the population, so they collect a random sample, say of 100 animals. Sampling at random is key to collecting data that represent the broad population, thereby avoiding biases in the parameter estimates. In this example, on a given day, for each animal, the researchers record the number of calls produced during daylight hours and during the night. Let us call the event, in which for a given animal there are more calls during the day than at night, a "success." If we assume animals operate independently, then the number of successes in the 100 animals provides information about the null hypothesis: the further from the expected number if there were no differences between night and day, the larger the evidence against $H_0$. We also assume that the probability of a success is constant and independent across trials and animals. Under $H_0$ we assume the probability of a success is $p = 0.5$. Under $H_0$, the number of successes has a binomial distribution with parameters $n$ (the sample size) and $p$. The corresponding probability mass function with $n = 100$ and $p = 0.5$ is illustrated in Fig. 9.9.

To test the null hypothesis, the researchers use the number of successes as a test statistic. The test statistic has information about the null hypothesis, and under the null hypothesis, we know the distribution of the test statistic. If call rates are on average the same during the night and day (i.e., $H_0$ is true), then we would expect that animals have a probability of 0.5 of producing more calls during the day than at night, and on average $T$ (number of successes) would equal 50 ($T = 50$).

Now imagine that the researchers observe $T = 46$. From Fig. 9.9, $T = 46$ is consistent with the null hypothesis, which we would not reject for the usual levels of statistical significance (see below for a more in-depth discussion of significance levels). On the contrary, consider the case of $T = 11$. This result would have been extremely unlikely under the null hypothesis, and we would be tempted to reject the null hypothesis, implying that differences between night and day might occur.

The example given here illustrates the rationale under NHST, the steps of which are: (1) define the hypothesis, (2) collect the data, (3) calculate a test statistic, with known distribution under $H_0$, (4) evaluate how likely (or unlikely) the data would be under the null hypothesis, and (5) if very unlikely, then reject the null hypothesis, but if not unlikely, do not reject it. Consequently, the trick is to put forward a null hypothesis under which the distribution of the test statistic can be evaluated to assess how likely the data are under the null hypothesis. Given the sampling uncertainty (i.e., not observing the entire population), we can make mistakes when making decisions about whether to reject the null hypothesis or not. The confusion matrix in Table 9.4 illustrates the possible outcomes of a decision.

The two wrong decisions we can make are to reject the null hypothesis when it is in fact true or

**Table 9.4** Confusion matrix showing the possible outcomes of a null hypothesis decision: correct decisions and Type I and Type II errors. Statistical tests usually require a significance level (i.e., Type I error rate), which defines the probability of being wrong if the null hypothesis is true

| | | Decision on null hypothesis | |
|---|---|---|---|
| | | Do not reject | Reject |
| Reality | $H_0$ true | Correct decision | Type I error |
| | $H_0$ false | Type II error | Correct decision |

to not reject it when it is false. The former is known as a Type I error (i.e., an incorrect rejection, sometimes referred to as a false-positive) and the latter a Type II error (i.e., failing to find a real effect, sometimes referred to as a false-negative). In general, it is believed that Type I error is what we should guard against, with the logic illustrated here as analogous to the legal system: *It is better to have a guilty defendant not convicted than to have an innocent defendant sent to death*. We note, however, that depending on the problem at hand, a Type II error could have a greater consequence than a Type I error. To illustrate this, imagine that you are testing whether the size of a population has decreased below a critical threshold that requires an action for it to not go extinct. If you do not reject the null hypothesis (i.e., that the population size has not changed) but it is false, you might miss the opportunity to take action and prevent the population's extinction. Alternatively, if you mistakenly take action to protect the population while it is in fact above the minimum threshold, you might waste money but any risk of detrimental population consequences is eliminated. So, while many textbooks may allude to the importance of safeguarding against Type I error, the error type that should be of most concern is likely to be study-specific. The usual advice applies: Do not use cookbook recipes, rather think about your study. The allowable Type I error can typically be specified with a critical significance level value (defined below). Estimation of Type II errors typically requires another step, called a power analysis (see Ellis 2010 for a textbook on power analyses).

In practice, the amount of evidence against the null hypothesis required in a study is given by setting a threshold based on how unlikely the observed data would have to be under the null hypothesis before it is rejected. Alternatively, we can compute the probability of, given the null hypothesis is true, observing a value for the test statistic that is as or even more extreme than the observed value. This probability value is commonly referred to as the *p*-value. In the above example, assuming a two-tailed test, the *p*-value associated with $T = 46$ or $T = 11$ would be 0.484 and ~0, respectively. This would lead us not to reject the null hypothesis in the first case, but to reject it in the second case. Note that a common error is to confuse the *p*-value with the probability of the null hypothesis being true or the alternative being false. Researchers should take care in their interpretation of *p*-values to ensure they are accurate.

The predefined probability threshold below which we are willing to reject the null hypothesis is called the significance level (typically designated as α). A typical value for the significance level is 5%, with tests having *p*-values lower than 0.05 often being reported as statistically significant. This value has become widely used; however, it should be noted explicitly that there is nothing special about a 5% significance level. While using this threshold has been extremely useful in practice, there is arguably no other concept in statistics that has received more criticism. The abuse of the 5% significance level by blindly using it is among the most common criticisms of the *p*-value and hypothesis testing (Nuzzo 2014; Yoccoz 1991; Beninger et al. 2012). Using common sense is fundamental in selecting significance levels. It is intuitively sensible that it cannot be sound science to blindly claim a result to be significant if $p = 0.049$ but not significant if $p = 0.051$. Ultimately, researchers need to think carefully about the cost of errors they can incur and define suitable significance levels accordingly. The focus should arguably

be on reporting confidence intervals and assessing the biological importance of reported effects, not on claims of statistical significance that are often not more than statements about sample size. Given a large enough sample size, even the smallest difference will become statistically significant. Therefore, it is perhaps not surprising that a common pitfall for researchers, and equally as or arguably more important than evaluating statistical significance, is failure to consider a result's biological significance. Imagine two populations of a whale species that produce the same stereotyped calls. Let us say animals in population A produced calls at a mean rate of 22.7 per hour and in population B at 22.6 calls per hour, and that these are significantly different statistically. Is this result meaningful biologically? In other words, is the effect size of a magnitude that we care about? In most cases, almost certainly not. Therefore, a researcher should have a good understanding a priori of the magnitude of the effect that is biologically relevant. Researchers undertaking studies with large sample sizes having the power to detect very small effect sizes can fall into the trap of reporting results as important based on statistical significance instead of on effect size and significance together. Conversely, studies having a large probability of incurring Type II errors (also known as low power, i.e., having a low probability of correctly rejecting the null hypothesis when it is false) due to a small sample size may only be able to detect very large effect sizes and miss smaller ones that are biologically important. The effect size that is meaningful in a study, thus, needs to inform the experimental design to ensure a sufficiently large sample is collected before the study commences.

While NHST and *p*-values can provide valuable tools to bioacousticians, it is not amiss for researchers to be well aware of the lively discussion on their misuse, drawbacks, and limitations. Nuzzo (2014) provides an introduction to this discussion, Yoccoz (1991) provides a classical critical review regarding their use in biology and ecology, and Beninger et al. (2012) frame the problem in the wider context of statistics in (marine) ecology. An entire *Forum* section in the journal *Ecology* has been dedicated to the topic in recent years, and Ellison et al. (2014) show that while having been discussed and revisited many times in recent years, the discussion about their use is alive and kicking!

Having said this, a wide range of NHSTs have been developed over many decades to accommodate a range of questions and data types. Traditionally, many of these have been described as either "parametric tests" or "non-parametric tests," with parametric tests often assuming samples arise from Gaussian distributions and non-parametric tests are often used for categorical or continuous data that do not fit assumptions of parametric tests. While we urge the reader to be cautious about blindly using such tests and be aware of their limitations, we feel we must discuss them since this is how statistics is presented in most undergraduate and postgraduate courses aimed at the applied sciences, biology and ecology included. As examples, tests commonly referred to as parametric include the z-test (for testing a sample mean), t-test (for comparing the means of two groups), and analysis of variance or ANOVA (used for comparing two or more groups). Common non-parametric alternatives to the t-test and the (one-way) ANOVA are the Mann–Whitney U and Kruskal–Wallis tests, respectively. The tests referred to here are only a few of the vast range available, and readers will not find it difficult to find a plethora of textbooks describing them. Note that these tests have been used widely in past decades and continue to be used in current research. Today, however, with improved knowledge of limitations of these tests, they are losing their appeal (see e.g., Touchon and McCoy 2016). In general, they are no longer the standard go-to for particular types of problems as they have been superseded by more robust approaches. With advances in statistics, a wide range of readily available modeling approaches has been developed that more than accommodate data that would have traditionally been analyzed using non-parametric tests (see Sect. 9.5.3 for an overview). Note that while many disciplines are guided by traditional "parametric" and "non-

parametric" classifications, where parametric would often be associated exclusively with the Gaussian distribution, modern approaches in statistical ecology using regression models are generally not said to be parametric or non-parametric; rather, they tend to be referred to based on the data distributions for which they are suited, such as a Poisson or gamma regression (see below for more on these).

### 9.5.3 Explanatory and Predictive Research Questions

Explanatory and predictive studies have questions requiring a response variable to be described as a function of a set of independent variables. Arguably, the majority of the models used by ecologists to answer this type of question are some kind of regression model. However, these models come in many forms. This section aims to introduce the reader to different types of regression models. We note upfront that model selection and validation, and inference from selected models, are fundamental aspects of these analyses and are only very briefly mentioned in Sect. 9.5.3.1. Relevant yet accessible books with plenty of practical examples addressing these steps include Zuur et al. (2007) and Zuur et al. (2009).

Historically, linear regression models (in which the errors are assumed to follow a Gaussian distribution) were the only tools available to answer this type of question. When the only tool you have is a hammer, all your problems begin to look like nails. With a Gaussian error distribution assumption, the only analytical options are simple linear regression models of the type given in Eq. (9.1) or linear regression models with several predictors (i.e., multiple regression). There are many special cases of such linear normal regression models including the independent sample t-test, ANOVA (i.e., analysis of variance for multiple sample mean comparison), ANCOVA (i.e., analysis of covariance for regressing a continuous response variable on a factor and a continuous covariate), and MANOVA or MANCOVA (i.e., multivariate extensions of the former methods). Note that these approaches have additional assumptions, such as that of homogeneity of variances. Homogeneity of variance means that the variance for a response variable is assumed to be constant across values of the independent variable. Many datasets have been forced through these methods even when they were clearly not the right tool for the job. This included, for example, transforming the response variable (e.g., by applying a log function to it) until Gaussian distributional assumptions were met to a reasonable extent. But even then, often a method's assumptions were not met. For instance, there is no transformation that will turn a discrete count into a continuous variable. For an interesting presentation about why not to log-transform data, see O'Hara and Kotze (2010). Nonetheless, sometimes processes might have properties that make a log-transformation of the data sensible and useful (e.g., Kerkhoff and Enquist 2009). While transforming data to fulfill methods' assumptions has been acceptable in the past given a lack of accessible alternative methods, this is often no longer the case, and successful ecologists need to have a few additional tools in their toolbox. The rule is one that practitioners do not enjoy: There is not a single rule that fits all questions and problems, we need to understand the problem to know how to model it. Sometimes it is even said that modeling is as much an art as it is a science. But like any good artist, you must master the techniques to use them correctly.

The next level of sophistication in regression models came with the advent of Generalized Linear Models (GLMs). GLMs allow for different types of response variable and some degree of non-linearity in the relationship between the response and explanatory variables. The relationship will still be linear at some level, but it might not be at the response level, it might only be linear at the level of the link function. What is the link function? It is a fundamental component of a GLM and is what allows responses to be constrained to a specific range of values. The link function, as its name implies, links the linear predictor and the response variable so that the model equation looks like:

$$g(E(Y)) = \alpha + X\beta, \tag{9.2}$$

where $g$ is the link function, $E(Y)$ is the expected value of the response variable, and as in simple linear regression (see Eq. 9.1), $\alpha$ is the intercept (a constant), $X$ is the predictor variable, and $\beta$ is the regression coefficient. For a vector of $n$ observations, the equation is in matrix form, where $\beta$ is a vector of parameters and $X$ is a matrix of predictor observations. The presence of a link function in Eq. (9.2) means that to obtain a prediction from this model, we need to apply the inverse of the link function to the linear predictors. As an example, consider a model with a log-link function. The inverse of the log is the exponent. This means that we need to exponentiate linear predictors to obtain the predicted value of $Y$ for the corresponding values. But then, this also means that, irrespective of the covariate values and the coefficients estimated, the prediction will be positive (because the exponent of any number is positive). Some link functions allow values predicted for the response variable to be constrained (limited) to between 0 and 1, further increasing the range of modeling possibilities to include binary responses (e.g., presence/absence) or proportions. For instance, binary response variables like presence/absence are modeled using a binomial GLM, with logistic regression being a special case of a binomial GLM, where the link function is the logit function. Count data can be modeled using a Poisson GLM. The Poisson distribution is quite inflexible, however, because as noted above, it assumes that the mean and the variance are the same. Quite often, biological data are overdispersed, meaning that the variance is greater than the mean. For such count data, a quasi-Poisson or negative binomial response is often a second natural choice as it allows the variance to be greater than the mean. Finally, we could also consider other less commonly used, but equally useful, GLMs: (1) multinomial regression when the response can take one of several categorical outcomes, (2) gamma regression where the response is strictly positive, and (3) beta regression when the response is a probability or a proportion.

While GLMs allow added flexibility to standard linear regression as a result of the link function, if the relationship between the response and the predictors is highly non-linear (i.e., cannot be assumed linear even on the link function scale), then a GLM will not be adequate. This is where we need to bring non-linear functions into play, and perhaps the most widely used non-linear approach is the Generalized Additive Model (GAM). GAMs also consider a link function to allow different distributions for the response variable (as in GLMs), but we now have the response being a function of smooth functions of the predictors. In a univariate case, the model equation looks like:

$$g(E(Y)) = \alpha + f(x), \tag{9.3}$$

where $g$ is the link function, $E(Y)$ is the expected value of the response variable, $\alpha$ is the intercept, $x$ is the predictor variable, and $f$ is a function such as a polynomial or spline. The polynomial or spline applies a smooth, curved-type function to the variable.

All the models described so far, be it a simple linear model (LM), a GLM, or a GAM, include only independent variables that are considered to be fixed effects. However, sometimes the inclusion of random effects might be necessary. A random effect is useful when we have observed a (random) subset of a larger population of possible values for a covariate. For example, a study may be interested in identifying responses of bats from a certain population before, during, and after exposure to high-frequency sound. The individual bats, whose responses were measured before, during, and after exposure, are a random effect. Random effects can be incorporated into a range of linear regression type models. For instance, Generalized Linear Mixed Models (GLMM) and Generalized Additive Mixed Models (GAMM) are GLMs and GAMs that incorporate both fixed and random effects. The reader is referred to Harrison et al. (2018) for an overview of mixed models in ecology, Pedersen et al. (2019) for non-linear models including mixed effects, and Nakagawa and Schielzeth (2010) for a review of the general issue of dealing with repeated

measurements sharing a correlation structure in biological studies.

Despite these advances, some data still do not fit the distributional requirements of GLMs and GAMs. Generalized Estimating Equations (GEEs) have been introduced recently, and hence they might still be considered in their infancy, but they are showing promising results. GEEs generalize GLMs and GAMs even further by not requiring that the response variable come from a particular family of distributions. GEEs simply impose a relationship between the mean and variance of the response. These models also allow a wide range of correlation structures to be imposed on the data, making them quite appealing when there are many observations clustered inside a few individuals. GEEs are marginal models in that the focus of inference is on the population average, and we are not so interested in the responses at the individual level. GEEs are quite specialized, and the reader is referred to Zuur et al. (2009, Chap. 12) for an introduction.

In addition to the somewhat "general" regression models above, there is a range of specialized regression models that are worth considering in certain biological questions. For instance, we have mentioned the problem of overdispersion. Often with biological data, we have very special cases of overdispersion in which there is an excess of zeroes. For example, consider you are trying to model the number of echolocation clicks a sperm whale produces per second as a function of depth, time of day, and sex. There are (at least) two reasons for there being zero clicks in a given second. A whale is in a silent state when recorded and many zeroes occur in successive seconds, or the whale is in a click-producing state but does not produce a click in the given second recorded. The regression models discussed above will likely fail to produce reasonable answers because the excess zeroes from the silent periods (potentially not explained by the covariates; i.e., not dependent on sex, depth, or time of day) cannot be accommodated. Under such a scenario, hurdle models or zero-inflated models might come in handy. While these are advanced methods and more difficult to implement and evaluate, they are worth knowing about. The reader is referred to Martin et al. (2005) for a gentle introduction to the topic with ecological examples.

Truncated regression is another special case of regression under which some values of the response variable cannot be observed. An example is modeling animal group sizes as a function of their acoustic footprint (e.g., the number of sounds produced by a group that are detected per minute). Now that you know about GLMs, your first thought might be to consider a Poisson or negative binomial GLM, with group size as the response variable and numbers of sounds detected as the predictor. However, in modeling this, you soon face a problem: You fit your model and make some predictions, one of which is a group size of zero! What does this mean? Nothing really, it is what we call an inadmissible estimate and a clear sign that something is not adequate. Under such a case, you might want to try a zero-truncated regression, which is essentially a GLM for which zeroes cannot be observed. Chapter 11 in Zuur et al. (2009) explores both zero-inflated and zero-truncated models.

Survival models are regression techniques that deal with a special type of response variable: the time up to an event. While these types of models were developed to model survival of animals, plants, and people, they can be used in any scenario where observations might be censored. Censored data result when we do not know the real value of the response variable but know it is at least above or below some limit or within some interval; say because we observe an animal is dead at a given time, and/or we know it was alive at a different time. For example in a bioacoustic study, a researcher may wish to model the time animals take to produce their first acoustic cue, and animals are observed for 5 min each. However, we do not know when an animal produced a cue before observations began (i.e., left censoring). In addition, an animal might not produce any cues during the 5 min, or the animal might leave the study area before the 5 min elapse (i.e., right censoring). Finally, if we recorded only which minute, but not the actual second a sound was produced, we would only know that the event occurred sometime within the interval of that minute. These are interval

**Table 9.5** Description of some commonly used models to test the association between multiple explanatory variables and a response variable

| Model type | Use |
|---|---|
| Generalized Linear Modeling (GLM) | Allows different distributions for the response variable and some degree of non-linearity in the relationship between response and explanatory variables |
| Generalized Linear Mixed Effects Modeling (GLMM) | An extension of GLM for use with random effects (e.g., repeated measures of subjects) |
| Generalized Additive Modeling (GAM) | Allows different distributions for the response variable (as in GLMs) modeled as a function of smoothed predictors |
| Generalized Additive Mixed Effects Modelling (GAMM) | An extension of GAM for use with random effects (e.g., repeated measures of subjects) |
| Generalized Estimating Equations (GEE) | Do not require the response variable to come from a particular family of distributions, and allows correlation structures in the data to be accounted for |

censored data. While a somewhat contrived example, this allows us to introduce the different kinds of censoring that are common in survival analysis.

Generalized Least Squares (GLS) is a regression approach that might be used when we want to relax the usual assumption of homogeneous residual variance by modeling the variance as a function of covariates. Zuur et al. (2009, Chap. 4) provide examples of the use of GLS and Reyier et al. (2014) give an acoustics application of GLS. Another perhaps more specialized use of such a regression technique is when we want to consider a general non-linear model with a specific form to relate a response variable with covariates. Then we might still want to find the parameters of the model that best fit the data. A way to do so is, akin to what might happen if one considers a straight line, to find the parameter values that minimize the sum of the squares of the residuals (i.e., the difference between the observations and the model). In a simple regression context, the model produces the fitted line, while in a generalized least squares context, the model is any function in which we might be interested. For example, if you want to determine the propagation loss ($PL$) for a sound that has traveled from the source to the receiver, and you expect it is proportional to $\log(r)$, where $r$ is the range, then your model is $PL = K \log (r)$. Based on measurements of received levels of sounds with known source level, you may apply a GLS regression to estimate the value of $K$ that best fits your data. If $K$ is close to 10, then your environment supports cylindrical spreading, if it is close to 20, then sound is predicted to spread spherically (see Chaps. 5 and 6 on sound propagation in air and under water, respectively).

All the models described so far do not consider predictor variables that are in hierarchies. Hierarchical data occur when variables are nested within each other (i.e., organized into levels). For example, individuals from different resident populations can be said to be nested within subpopulations. In turn, subpopulations can be nested within populations. Hierarchical modeling (also known as multilevel modeling) is used when inferences need to be drawn for population means at specified levels and is useful for fitting models to data obtained from complex, multilevel survey designs. For example, a study may evaluate vocal complexity of elephants at the population, sub-population, and resident population levels. Here, we do not discuss these methods further. Rather, we refer the reader to Cressie et al. (2009) and Royle and Dorazio (2008) for descriptions of these methods, including their strengths and limitations.

Given the large range of models available (a taste of which has been described above), what should aspiring ecologists today have in their statistical regression toolbox? We propose that a bare minimum is an understanding of the structure, implementation, outputs, and interpretation of GLMs, GLMMs, GAMs, and GAMMs (Table 9.5). Parameter estimates and significance tests resulting in $p$-values are common outputs of software capable of fitting GLMs, GLMMs, GAMs, GAMMs, and GEEs. For a practical guide to applying these in behavioral and

ecological studies, see Zuur et al. (2009). O'Hara (2009) and Bolker et al. (2009) provide good introductions to GLMMs for ecologists, and the books by Zuur et al. (2007, 2009) provide information to implement and interpret GLMMs. For GAMs, the book by Wood (2006) is a standard reference, and Zuur et al. (2009) has worked-out examples in the software R.

Most of the models described in this section can be implemented in a frequentist framework, for instance using maximum likelihood or restricted maximum likelihood estimation. Nonetheless, for more complex models such as those including (often complex) spatial and temporal covariates (i.e., spatio-temporal models), Bayesian implementations are gaining ground. For instance, GLMs and GLMMs are fitted via maximum likelihood, or Markov Chain Monte Carlo (MCMC). MCMCs are Bayesian iterative solutions and are described in Gamerman (1997), Brémaud (1999), Draper (2000), and Link (2002). With advances of widely available implementations, users might even be using Bayesian approaches without realizing it. An example is the Integrated Nested Laplace Approximation (INLA) implemented via R-INLA (www.r-inla.org) and its derivatives that allow fitting complex spatio-temporal models without the Bayesian framework being obvious (by not requiring priors to be explicitly defined). The philosophical nuances of which framework might be more adequate under given settings, however, are beyond what we hope to discuss in this chapter.

#### 9.5.3.1 Model Validation, Selection, and Averaging

Depending upon whether modeling is undertaken for explanatory or predictive purposes, approaches for model validation and selection may differ (Shmueli 2010). Validation means that the model has been demonstrated to have satisfactory accuracy for its intended use (Rykiel Jr 1996). Validation in explanatory modeling commonly takes the form of goodness-of-fit and residual diagnostics. Goodness-of-fit tests evaluate how well-observed values agree with those expected under the statistical model (Maydeu-Olivares and Garcia-Forero 2010), while residual diagnostics determine whether residuals fit the assumption of being effectively random (see Zuur et al. 2009 for common examples in ecology). Checking for multi-collinearity (i.e., collinearity between two or more covariates) is also standard for explanatory modeling, while it is close to irrelevant for predictive modeling (see Shmueli 2010 for detailed discussion). In contrast to explanatory modeling, model validation in predictive modeling is focused on evaluating the model's ability to generalize and predict new data. Validation commonly is undertaken using approaches such as cross-validation. In cross-validation, the model's ability to accurately predict a new data set is assessed after calibrating it with a training dataset (Shmueli 2010; Cawley and Talbot 2010).

Once a set of models have been validated, the best candidate model is selected (though model validation and selection can often be an iterative process). Approaches to model selection, again, depend upon whether modeling has an explanatory or predictive goal. In explanatory modeling, the explanatory power of nested candidate models is commonly compared with a step-wise approach using significance testing (e.g., using an $F$-test). Here a nested model refers to one composed of subsets of covariates of another candidate model. Caution should be taken, however, as researchers may be inclined to remove covariates that are not significant, even when there is a strong theoretical justification for retaining them since they are relevant in the models, regardless of whether they are significant or not (Shmueli 2010). For example, a covariate representing the age class of a sparrow in a study assessing the influence of predator presence on sparrow vocal behavior may be of theoretical importance in the model. Model selection in predictive modeling commonly involves a priori specification of candidate models and selecting the best model based on the smallest possible number of parameters that adequately represent the data (i.e., the principle of parsimony). The simpler a model is, the more it can be generalized, while more complex models (containing more parameters) are more specific to the data used to fit the model. Consequently,

criteria for model selection have been developed that essentially maximize the likelihood while penalizing for the number of parameters included. The Akaike's Information Criterion (AIC; see Akaike 1974) and Bayesian Information Criterion (BIC) currently are the most commonly used, among a range of others available. They are widely used for comparing nested and non-nested models (Burnham and Anderson 2002), although there is some discussion around suitability for use in non-nested models (see Ripley 2004). Resulting criteria such as AIC or BIC values for candidate models are then compared and the model yielding the lowest value is generally deemed to be preferred. Note that there is active research on the circumstances under which AIC, BIC, and the many other criteria available perform best, and whether they should be used together to inform model selection (Kuha 2004). An important take-home message is that model selection criteria such as AIC and BIC can only suggest a preferred model from those compared, even if they all perform poorly at the validation stage. In other words, the preferred model may still be a poorly fitting model, and therefore, selection criteria are only relative measures of model goodness-of-fit.

In predictive modeling, averaging over a range of plausible models has become widely used to reduce prediction error and improve model selection uncertainty. This is undertaken, for example, by computing a measure that ranks the set of plausible models according to their support by the data (e.g., Akaike weights), applying the weights to predictions from each model, and then computing the average. This provides weighted averaged predictions, with weights dependent on how much each model is supported by the data. There are many other methods for undertaking model averaging. Model averaging performance depends on each model's predictive bias and variance and covariance between models, among other things (see McElroy 2016 for complete discussion). In recent work, model averaging has been shown to be particularly useful when predictive errors of contributing model predictions are dominated by variance, and when covariance between models is low (McElroy 2016).

While a highly simplified overview of some tools available on the topic of model validation, selection, and averaging has been provided here, researchers should be familiar with them and access the latest literature to identify the appropriate approaches for their study.

### 9.5.4 The Future of Bioacoustical Analytical Approaches

In this chapter, we have only provided a flavor of common approaches used today and have not delved into the wide range of new developments being introduced into the discipline. Interdisciplinary research linking the fields of biology, ecology, and statistics has a long tradition of providing fertile ground for innovative statistical methods, with many methods having been developed when existing methods were not adequate to cope with new problems (Olivier et al. 2014). The current revolution in data acquisition systems (see Chap. 2), such as high-resolution sensors in animal-borne tags and increasing numbers of long-term passive acoustic deployments that lead to big data, is also likely to influence the next generation of statistical methods suited for ecological and acoustical analysis. Analysis of big data through increased computational capacity has already provided a range of new powerful tools to science.

As an example of such approaches, machine learning is rapidly gaining in popularity as it increasingly improves pattern recognition accuracy (Christin et al. 2019). Such methods can improve processing capacity in large datasets resulting from acoustic instrumentation. An example of more sophisticated analytical approaches is the growing use of hierarchical, state-space, and hidden process methods (e.g., Auger-Méthé et al. 2020 for an introduction to their application in ecology) that model underlying processes while accounting for biases and uncertainty. Advances in these approaches may improve our ability to predict future scenarios and implement intervention before a potentially

undesirable future scenario unfolds (see Cressie et al. 2009 for discussion).

We also suggest readers to be acquainted with the growing work being conducted in the area of statistical decision theory, which is concerned with making decisions by accounting for uncertainties involved in the decision process using statistical knowledge resulting from data collected. Rather than attempting to provide a general review of the large field of decision theory here, we refer the reader to an introduction in its application to ecology by Williams and Hooten (2016), which will introduce the reader to a range of other resources on the topic.

Because the advancement of these and many other methods are continually evolving, researchers are encouraged to keep well-informed of current developments appearing in methods-based scientific journals, such as *Methods in Ecology and Evolution*.

## 9.6 Examples in Bioacoustics

The wide range of quantitative approaches introduced above can be used to analyze bioacoustical data to answer research questions ranging from understanding natural vocal behavior to activity patterns, community and conservation ecology, habitat use, species diversity, distribution, occupancy, density and abundance, and anthropogenic impacts (among many others). Faunal groups that have been the subject of bioacoustics research include invertebrates, anurans (i.e., frogs and toads), fish, birds, bats, other terrestrial mammals, and marine mammals, but many others could be considered. As long as sound is produced, it could be used as a source of information. A recent review documented 460 peer-reviewed published papers on passive acoustic monitoring in terrestrial habitats alone, with bats (50% of papers) and activity patterns (24%) dominating (Moreria Sugai et al. 2018). Marine mammals feature prominently in bioacoustic research as water is a highly conducive medium for sound to travel through, and visual observations can prove comparatively expensive for limited returns on detections. Rather than reviewing analytical approaches across the hundreds of existing bioacoustics studies, we have selected two recent studies as examples, and discuss the rationale for the particular analytical approaches taken. The research topics in the example studies are exploring temporal changes in call frequency and using acoustic data for abundance and density estimation.

### 9.6.1 Temporal Changes in Call Frequency

As indicated previously, due to ever-increasing computing power and storage and technological advances in acoustic equipment, acoustic studies can provide extremely long-term datasets. These datasets allow us to explore changes to calling behavior on a scale that, until recently, would have been very difficult. A recent example is illustrated in Miksis-Olds et al. (2018) where the frequency content of a type of blue whale song recorded primarily in the Indian Ocean was investigated. The song type is attributed to a pygmy blue whale subspecies (*Balaenoptera musculus indica,* Committee on Taxonomy 2021) that appears to be resident in the northern Indian Ocean. The song type has three distinct units, and this analysis focused on the ~60-Hz component of Unit 2, a frequency-modulated upsweep, and Unit 3, a ~100-Hz tonal downsweep. A decade of data from the Indian Ocean Comprehensive Nuclear-Test-Ban Treaty International Monitoring Station (CTBTO IMS) at Diego Garcia was analyzed (2002–2013). Ambient noise was also analyzed, but we do not focus on that part of the study here.

Power spectral densities (PSD) were computed for 2-h sections of data, which could be used to detect peaks in the frequency bands of interest (approximately 56–63 Hz for the 60-Hz component of Unit 2, and 107–100 Hz for Unit 3), using a 3-dB signal-to-noise threshold. The paper shows a figure of number of hours with vocal presence detected each week, for each year (Fig. 9.3 in Miksis-Olds et al. 2018), highlighting the importance of producing exploratory plots; in this case, the variability in the data is made clear.

The average over each week, across years, was used to identify weeks with peak average vocal presence. Weeks 21 and 22 were those with peak average vocal presence and data from these weeks were investigated further. The frequency peaks from the PSDs from these weeks across all years were measured. A linear regression model was fitted to the week 21 and 22 frequency peak measurements from all years. The response variable was frequency, and year and song unit were explanatory variables. Song unit was included in the model as a factor variable. An interaction was also included between year and song unit, which was used to investigate whether the rate of any frequency change over time differed between the two song units. Model assumptions (linearity, constant error variance, error independence, and normality) were all assessed using diagnostic plots and relevant hypothesis tests, and all model assumptions were met.

The linear model results are depicted in Fig. 9.10. The figure shows all weekly data plotted (blue dots) with the modeled 21–22 week data highlighted in red for both song units. Again, the utility of plotting data is clear here: the decline in frequency is evident, with an apparent difference in rate of decline between the two units. The linear model results confirmed the frequency decline; the frequency of the ~60-Hz Unit 2 decreased at a rate of 0.18 Hz/year, while the frequency of Unit 3 decreased at 0.54 Hz/year. The interaction term was selected during model selection (using an F-test), which confirmed that the rates of frequency decline were indeed different between the two units.

This analysis shows that simple regression analyses can be very effective in confirming patterns observed in exploratory data plots. We note here that the regression analysis in the paper focused on data from weeks 21 and 22 to be comparable with methods from a similar study (Gavrilov et al. 2012). However, frequency measurements were taken across all weeks of each year (as shown in Fig. 9.10), which could also be used in a regression model. In addition, it is common for bioacoustical analyses to have several natural extensions. In this case, relaxing the Gaussian assumption could be considered via a Generalized Linear Model, or non-linear patterns in the frequency decline could be explored using a Generalized Additive Model.

### 9.6.2 Abundance and Density Estimation

The estimation of animal population size (abundance) and the number of animals in a given area (density) are metrics that are very informative for management and conservation actions. There are several abundance and density estimation methods available (e.g., Borchers et al. 2002); popular methods include mark-recapture and distance sampling. Such methods are known as absolute abundance or density estimation methods, as the methods estimate the total number of animals (in a defined area, for density estimates), including animals missed by a survey. Common reasons why animals are not detected during a survey is that they may be too far away, and/or detection is made difficult by environmental conditions (e.g., rough seas may prevent marine mammal sightings at sea unless the animals are very close, or windy conditions may mask the sounds of singing birds in recordings). The probability of detecting an animal is a key parameter in absolute abundance and density estimation methods, and accounts (in part) for undetected animals during a survey.

Acoustic data are increasingly being used for absolute abundance and density estimation, both in terrestrial and marine environments (e.g., Marques et al. 2013; Stevenson et al. 2015). Here we discuss a density estimation analysis for Blainville's beaked whales (*Mesoplodon densirostris*) from seafloor-moored hydrophone data recorded in the Bahamas (Marques et al. 2009). The analysis involved several of the concepts we have discussed throughout the chapter, which we highlight here.

The paper begins by introducing the density estimation equation (i.e., the estimator; see Sect. 9.4.2). The equation contains several parameters to be estimated, including the probability of detecting a beaked whale echolocation click on one of the seafloor-moored hydrophones. Survey

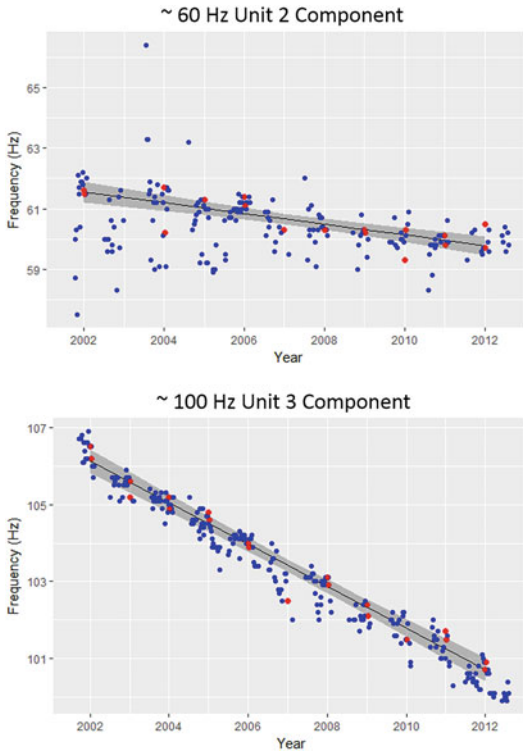



**Fig. 9.10** Peak frequency of Sri Lankan whale vocalizations determined from weekly PSD sound averages. The blue circles are the weekly peaks measured throughout the season when whales were vocally present. The trend line is related to the red circles that are peak frequency from weeks 21 and 22 of each year. The greyed regions designate the 95% confidence intervals for the trend. Reprinted with permission from Miksis-Olds et al. (2018). © Acoustical Society of America, 2018. All rights reserved

design and variance estimation of the parameters (including confidence intervals) are also discussed. A summary of methods to estimate the detection probability is given. Mark-recapture and distance sampling methods are commonly used approaches to estimate the detection probability, but Marques et al. (2009) needed an alternative method, given that the hydrophone recordings were not suitable for either mark-recapture, or distance sampling-based methods. Therefore, a trial-based detection probability estimation method was used. The specific trial-based method used in this study relied on auxiliary data from animals tagged with acoustic tags, which swam near the moored hydrophones. Clicks produced by the animals and recorded on the tags created "trials"; a successful trial was achieved if the same clicks recorded on tags of the tagged animal were detected on the moored hydrophones. In addition, the tag data provided the slant distance of each tagged animal from the moored hydrophones, as well as the animal's orientation toward, or away from, a given moored hydrophone. These data allowed detection probability to be modeled as a function of a whale's orientation and distance from the moored hydrophones using regression modeling. Specifically, a Generalized Additive Model (GAM) was used due to its flexibility in allowing non-linear relationships between the response and explanatory variables. The response variable was defined as the detection, or non-detection, of each click produced by the tagged animal on the moored hydrophones. The explanatory variables, or covariates, were (a) the horizontal off-axis angle (hoa) and (b) vertical off-axis angle (voa) of the tagged whale, with respect to a given moored hydrophone, and (c) the distance of the tagged whale from the hydrophone. A binomial distribution was assumed for the response variable due to the binary nature of the trial data (i.e., detected, or not detected) and a logistic link function was used in the GAM. Finally, to estimate the average detection probability (i.e., a single parameter value for the estimator), a Monte Carlo simulation was implemented where the dive profiles from the tags were randomly placed around virtual moored hydrophones. In the simulation, the slant range and orientation of the clicks from the dive profiles from the moored hydrophones could be calculated, and then these values could be used along with the GAM to predict the detection probability for each click in the simulation. The average of these predicted detection probabilities was used in the estimator. Two other parameters required for the estimator, the false-positive proportion and cue production rate, are discussed in the paper in detail, on which we do not focus here.

The results of the GAM are shown in Fig. 9.11. The modeled relationships between (a) detection probability and slant range, (b) vertical and horizontal off-axis angle and

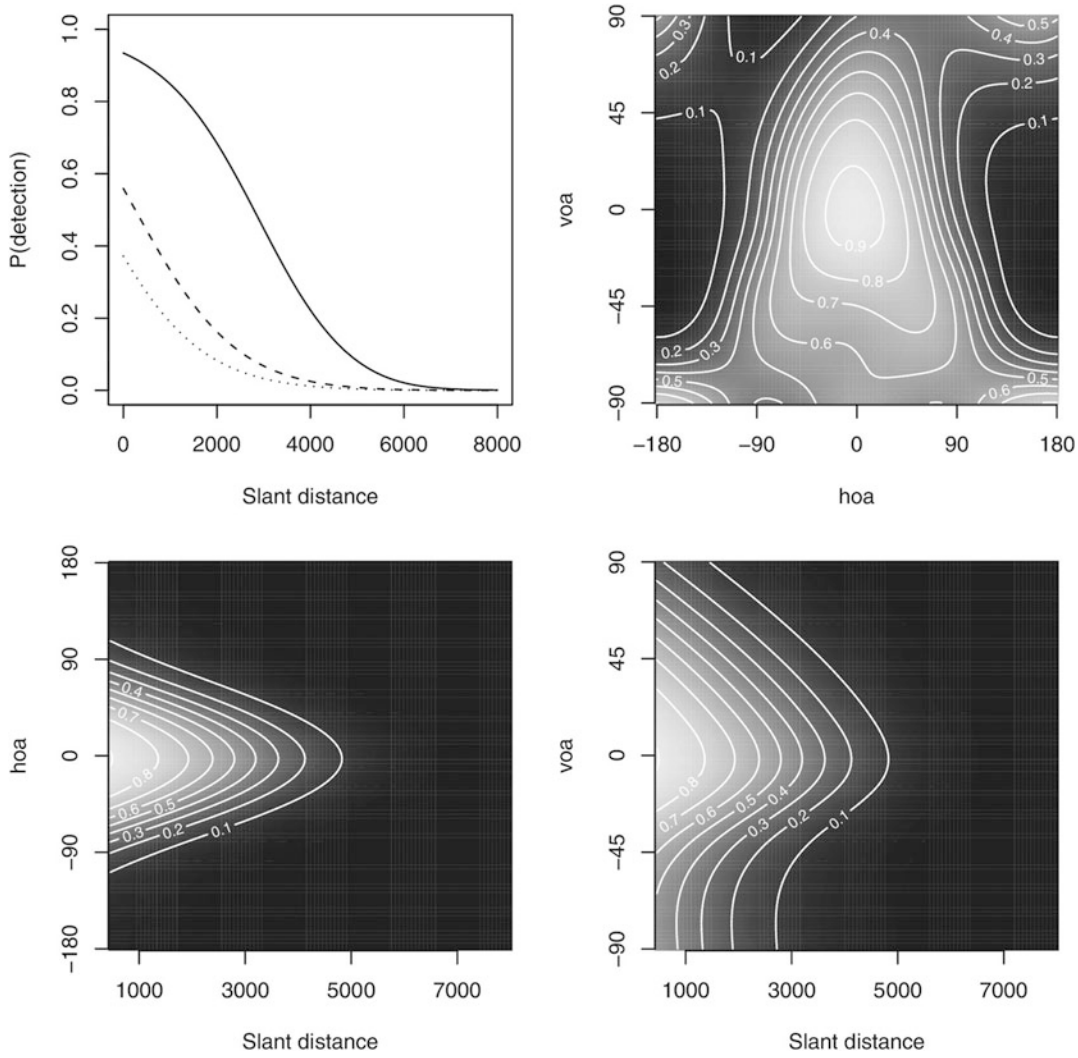



**Fig. 9.11** The estimated detection function. Plots (on the response scale) of the fitted smooths for a binomial GAM model with slant distance and a 2D smooth of *hoa* and *voa*. For the top left plot, the off-axis angles are fixed at 0, 45, and 90° (respectively the solid, dashed, and dotted lines). Remaining plots are two-dimensional representations of the smooths, where black and white represent respectively an estimated probability of detection of 0 and 1. Distance (top right panel) and angle not shown (bottom panels) are fixed respectively at 0 m and 0°. Reprinted with permission from Marques et al. (2009). © Acoustic Society of America, 2009. All rights reserved

detection probability, (c) horizontal off-axis angle and slant range, and (d) vertical off-axis angle and slant range are all depicted. The average detection probability of a beaked whale click within 8 km of a moored hydrophone was estimated to be 0.03 (i.e., if a beaked whale click was produced within 8 km of a moored hydrophone, the study estimated that there was, on average, a 3% chance of detecting that same click). The variance around the average was estimated using the bootstrap and presented as a coefficient of variation (CV, defined in Sect. 9.4.2) and was estimated to be 0.16, or 16% when expressed as a percentage. Finally, the estimator was used to estimate beaked whale density in the study area of either 25.3 (CV: 19.5%) or 22.5 (19.6%) animals per 1000 km$^2$,

depending on the false-positive proportion used (two estimates were produced using differing methods).

## 9.7 Software for Analyses

There are many standard, relatively easy-to-use software packages that require no (or very little) coding skills to carry out statistical analyses, including *SPSS* (IBM Corp., Armonk, NY, USA), *Statistica* (TIBCO Software, CA, USA), *Stata* (StataCorp, College Station, TX, USA), *Minitab* (Minitab Inc., State College, PA, USA), *Xlstat* (Addinsoft, Ile-de-France, France), and *SAS* (SAS Institute, Cary, NC, USA), among others. In the field of bioacoustics, it is common for acoustic data to be processed in *MATLAB* (The MathWorks Inc., Natick, MA, USA) due to its powerful signal processing package. *MATLAB* users may find that their workflow is streamlined by undertaking statistical analyses in the same software if all required tools are available.

For those planning, however, on undertaking analyses that draw from the most recent up-to-date developments in statistical ecology and require a highly flexible environment to do so, a free open-source software environment like R is recommended (R Core Team 2020). R is primarily used for statistical computing and production of graphics (though R's GIS, and even signal processing capabilities, are expanding). The software benefits from a large number of base and contributed packages that can easily be downloaded and an environment in which users may develop their own algorithms and packages. There are now many sources of instructional manuals and books guiding users on how to create high-quality data representations and run analyses in R, including Crawley (2013), Kerns (2010), Zuur et al. (2009), Bolker (2008), Lawson (2014), among many others. The *CRAN Task View: Analysis of Ecological and Environmental Data*[1] maintained by Gavin Simpson is an excellent resource for locating suitable packages for statistical analysis of biological data. R can be accessed and downloaded through a web browser[2] and for most users, we recommend a user-friendly GUI like *RStudio* (RStudio Team 2020[3]). *RStudio* is an integrated development environment for R that includes a console, an editor for code development and execution, and tools for plotting, debugging, tracking history, and managing the workspace. An interesting feature of R integrated with *RStudio* is the ability to adhere in a straightforward way to the concept of reproducible research via dynamic reports in *RMarkdown*. If the reader is new to the topic, we recommend the book by Xie et al. (2020).[4]

## 9.8 Summary

A key outcome of bioacoustics research is the production of new knowledge that informs conservation management. The knowledge produced needs to be reliable and easily understood, which is no trivial task given the complicated nature of animal behavior. The reality is that the phenomena from which we want to derive inferences are multifaceted, with many interconnecting attributes, and patterns and signals obscured by statistical noise (i.e., variability not associated with the conditions under investigation). Consequently, underlying mechanisms that explain the patterns we observe are not easily revealed.

Not only are animal behaviors occurring in a highly complex environment, but many challenges are presented in conducting the research itself. For instance, as researchers we are not easily able to avoid or reduce the statistical noise in the environment by controlling field conditions; and when we undertake experiments of animals in captivity to reduce noise in a laboratory, we cannot be sure that results are

[1] CRAN Task View: https://CRAN.R-project.org/view=Environmetrics; accessed 9 November 2020.

[2] R Core Team is accessible at https://www.r-project.org/; accessed 1 January 2020.

[3] *RStudio* is accessible at https://www.rstudio.com/products/RStudio/; accessed 9 November 2020.

[4] *RMarkdown*: The Definitive Guide by Xie Y, Allaire JJ, Grolemund G: https://bookdown.org/yihui/rmarkdown/; accessed 9 November 2020.

transferable to the wild. In addition, we introduce biases in our observations through our own subjective, non-random filters. Only by understanding these filters can we either eliminate or adjust biases to make reliable inferences about nature.

Quantitative skills, including survey design considerations, are therefore an essential part of a bioacoustician's toolkit and should be viewed just as essential as field skills and signal processing methods. These statistical methods are tools that enable the researcher to ask difficult but often important and exciting questions about their research topic.

However, given the complexity in nature, research design challenges, and the multidisciplinary nature of studying animal behavior through acoustics, it is not realistic to expect specialists in one field to become experts across multiple fields (i.e., behavior, ecology, bioacoustics, and statistics). What behaviorists and bioacousticians can aim for is to understand foundational statistical concepts, have a broad knowledge of the range of existing techniques available, and be able to identify critical pitfalls in survey design and data analyses. In addition, practitioners should be able to conduct a range of current standard analyses and know when to seek support for more sophisticated approaches.

It is our hope that through the introduction of basic statistical concepts in this chapter, readers can more confidently avoid design and analysis pitfalls and make the necessary considerations to select the most suitable approaches to successfully answer their research questions. We would like researchers to feel empowered to critically evaluate the transferability of standard practices across broader spectra of questions and identify inadequacies where they occur. Finally, and foremost, we hope that at the conclusion of this chapter, readers feel inspired to place greater focus on the biological significance of research outputs, using quantitative methods as a tool to support their conclusions.

We close this chapter by providing you, the reader, with our culinary rendition of the meaning of statistics: It is the science that uses data as its main ingredient, uncertainty as a key seasoning driving the final flavor of a meal, and guides the collection and mixing of the ingredients, through sampling, experimentation, and analysis. Taken together, hopefully, delicious scientific meals will result, by drawing meaningful and reliable inferences from data. Statistics is paramount for science in general, and bioacoustics is in that regard no exception.

**Acknowledgement** We thank Steve Buckland and Jay Barlow for their helpful comments prior to Springer's peer-review.

## References

Akaike H (1974) A new look at the statistical model identification. IEEE Trans Autom Control 19(6): 716–723

Auger-Méthé M, Newman K, Cole D, Empacher F, Gryba R, King AA, Leos-Barajas V, Flemming JM, Nielson A, Petris G (2020) An introduction to state-space modeling of ecological time series. arXiv preprint arXiv:2002.02001

Beninger PG, Boldina I, Katsanevakis S (2012) Strengthening statistical usage in marine ecology. J Exp Mar Biol Ecol 426:97–108

Bolker BM (2008) Ecological models and data in R. Princeton University Press, Princeton

Bolker BM, Brooks ME, Clark CJ, Geange SW, Poulsen JR, Stevens MHH, White JSS (2009) Generalized linear mixed models: a practical guide for ecology and evolution. Trends Ecol Evol 24(3):127–135

Borcard D, Gillet F, Legendre P (2011) Numerical ecology with R. Springer, New York

Borchers DL, Buckland ST, Zucchini W (2002) Estimating animal abundance. Springer, New York

Box GE (1976) Science and statistics. J Am Stat Assoc 71(356):791–799

Brémaud P (1999) Markov chains: Gibbs fields and Monte Carlo simulation. Springer, New York, pp 253–322

Brown A, Smith J, Salgado Kent C, Marley S, Allen S, Thiele BL, Erbe C, Chabanne D (2017). Relative abundance, population genetic structure and passive acoustic monitoring of Australian snubfin and humpback dolphins in regions within the Kimberley. https://doi.org/10.13140/RG.2.2.17354.06082

Burnham K, Anderson D (2002) Model selection and multimodel inference: a practical information-theoretic approach, 2nd edn. Springer, New York

Casella G, Berger RL (2002) Statistical inference. Duxbury, Belmont, CA

Cawley GC, Talbot NL (2010) On over-fitting in model selection and subsequent selection bias in performance evaluation. J Mach Learn Res 11:2079–2107

Christin S, Hervet É, Lecomte N (2019) Applications for deep learning in ecology. Methods Ecol Evol 10(10): 1632–1644
Cochran WG (1977) Sampling techniques, 3rd edn. Wiley, New York
Cohen J (1988) Statistical power analysis for the behavioral sciences, 2nd edn. L. Erlbaum Associates, Hillsdale, NJ
Cohen J (2013) Statistical power analysis for the behavioral sciences, 2nd edn. Routledge, New York
Committee on Taxonomy (2021) List of marine mammal species and subspecies. Society for Marine Mammalogy. www.marinemammalscience.org. Accessed 2 Sep 2021
Crawley MJ (2013) The R book, 2nd edn. Wiley, Hoboken, NJ
Cressie N, Calder CA, Clark JS, Hoef JMV, Wikle CK (2009) Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling. Ecol Appl 19(3):553–570
Dytham C (2011) Choosing and using statistics: a biologist's guide, 3rd edn. Wiley, Boca Raton, FL
Ellis PD (2010) The essential guide to effect sizes: statistical power, meta-analysis, and the interpretation of research results. Cambridge University Press, Cambridge
Ellison AM (2004) Bayesian inference in ecology. Ecol Lett 7(6):509–520
Ellison AM, Gotelli NJ, Inouye BD, Strong DR (2014) P values, hypothesis testing, and model selection: it's déjà vu all over again. Ecology 95(3):609–610
Everitt B, Hothorn T (2011) An introduction to applied multivariate analysis with R. Springer, New York
Fisher RA (1959) Statistical methods and scientific inference, 2nd ed. Oliver and Boyd, Edinburgh, UK
Ford ED (2000) Scientific method for ecological research. Cambridge University Press, Cambridge
Gamerman D (1997) Sampling from the posterior distribution in generalized linear mixed models. Stat Comput 7(1):57–68
Gavrilov AN, McCauley RD, Gedamke J (2012) Steady inter and intra-annual decreases in the vocalization frequency of Antarctic blue whales. J Acoust Soc Am 131:4476–4480. https://doi.org/10.1121/1.4707425
Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB (2013) Bayesian data analysis, 3rd edn. CRC Press, Boca Raton, FL
Harrison XA, Donaldson L, Correa-Cano ME, Evans J, Fisher DN, Goodwin CE, Robinson BS, Hodgson DJ, Inger R (2018) A brief introduction to mixed effects modelling and multi-model inference in ecology. PeerJ 6:e4794
Hilborn R, Mangel M (1997) The ecological detective: confronting models with data, vol 28. Princeton University Press, Princeton
Kerkhoff AJ, Enquist BJ (2009) Multiplicative by nature: why logarithmic transformation is necessary in allometry. J Theor Biol 257(3):519–521
Kerns GJ (2010) Introduction to probability and statistics using R, 1st edn. G. Jay Kerns, Youngstown
Kuha J (2004) AIC and BIC: comparisons of assumptions and performance. Sociol Methods Res 33(2):188–229
Lawson J (2014) Design and analysis of experiments with R, vol 115. CRC Press, Boca Raton, FL
Leek JT, Peng RD (2015) What is the question? Science 347(6228):1314–1315
Link RF (2002) Principal applications of Bayesian methods in actuarial science: a perspective. North Am Actuarial J 6(2):129
Manly BFJ (2007) Randomization, bootstrap and Monte Carlo methods in biology. CRC Press, Boca Raton, FL
Manly BF, Alberto JAN (2014) Introduction to ecological sampling. CRC Press, Boca Raton, FL
Marques TA, Thomas L, Ward J, Dimarzio N, Tyack PL (2009) Estimating cetacean population density using fixed passive acoustic sensors: an example with Blainville's beaked whales. J Acoust Soc Am 125(4): 1982–1994. https://doi.org/10.1121/1.3089590
Marques TA, Thomas L, Martin SW, Mellinger DK, Ward JA, Moretti DJ, Harris D, Tyack PL (2013) Estimating animal population density using passive acoustics. Biol Rev 88(2):287–309
Martin TG, Wintle BA, Rhodes JR, Kuhnert PM, Field SA, Low-Choy SJ, Tyre AJ, Possingham HP (2005) Zero tolerance ecology: improving ecological inference by modelling the source of zero observations. Ecol Lett 8(11):1235–1246
Matthiopoulos J (2010) How to be a quantitative ecologist. Wiley, Hoboken, NJ
Maydeu-Olivares A, Garcia-Forero C (2010) Goodness-of-fit testing. Int Encycl Educ 7(1):190–196
McCarthy MA (2007) Bayesian methods for ecology. Cambridge University Press, Cambridge
McElreath R (2020) Statistical rethinking: a Bayesian course with examples in R and Stan. CRC Press, Boca Raton, FL
McElroy TS (2016) Nonnested model comparisons for time series. Biometrika 103(4):905–914
McGarigal K, Cushman SA, Stafford S (2000) Multivariate statistics for wildlife and ecology research, 1st edn. Springer, New York
Miksis-Olds JL, Nieukirk SL, Harris DV (2018) Two unit analysis of Sri Lankan pygmy blue whale song over a decade. J Acoust Soc Am 144(6):3618–3626
Moreria Sugai L, Freire Silva T, Wagner Ribeiro J, Llusia D (2018) Terrestrial passive acoustic monitoring: review and perspectives. Bioscience 69. https://doi.org/10.1093/biosci/biy147
Nakagawa S, Schielzeth H (2010) Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. Biol Rev 85:935–956
Nuzzo R (2014) Scientific method: statistical errors. Nat News 506(7487):150
O'Hara RB (2009) How to make models add up—a primer on GLMMs. Annales Zoologici Fennici, BioOne, pp 124–137
O'Hara R, Kotze J (2010) Do not log-transform count data. Nat Precedings 1:118–122
Ortega A, Navarrete G (2017) Bayesian hypothesis testing: an alternative to null hypothesis significance

testing (NHST). In: Psychology and social sciences. IntechOpen
Paiva EG, Salgado Kent CP, Gagnon MM, McCauley R, Finn H (2015) Reduced detection of Indo-Pacific bottlenose dolphins (*Tursiops aduncus*) in an inner harbour channel during pile driving activities. Aquat Mamm 41(4):455–468
Panzeri S, Magri C, Carraro L (2008) Sampling bias. Scholarpedia 3(9):4258
Pedersen EJ, Miller DL, Simpson GL, Ross N (2019) Hierarchical generalized additive models in ecology: an introduction with mgcv. PeerJ 7:e6876
Quinn GP, Keough MJ (2002) Experimental design and data analysis for biologists. Cambridge University Press, Cambridge
Rahlf T (2019) Data visualisation with R: 111 examples. Springer Nature, New York
Reyier EA, Franks BR, Chapman DD, Scheidt DM, Stolen ED, Gruber SH (2014) Regional-scale migrations and habitat use of juvenile lemon sharks (*Negaprion brevirostris*) in the US South Atlantic. PLoS One 9(2):e88470
Ripley BD (2004) Selecting amongst large classes of models. In: Methods and models in statistics, In Honour of Professor John Nelder, FRS. World Scientific, New York, pp 155–170
Royle JA, Dorazio RM (2008) Hierarchical modeling and inference in ecology: the analysis of data from populations, metapopulations and communities. Elsevier, New York
Rykiel EJ Jr (1996) Testing ecological models: the meaning of validation. Ecol Model 90:229–244
Salkind NJ (2010) Encyclopedia of research design, vol 1. Sage, Thousand Oaks, CA
Shmueli G (2010) To explain or to predict? Stat Sci 25(3): 289–310
Stauffer HB (2007) Contemporary Bayesian and frequentist statistical research methods for natural resource scientists. Wiley, Hoboken, NJ
Stevenson BC, Borchers DL, Altwegg R, Swift RJ, Gillespie DM, Measey GJ (2015) A general framework for animal density estimation from acoustic detections across a fixed microphone array. Methods Ecol Evol 6(1):38–48. https://doi.org/10.1111/2041-210X.12291
RStudio Team (2020) RStudio: integrated development for R. RStudio, PBC, Boston, MA. http://www.rstudio.com/
R Core Team (2020) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. https://www.R-project.org/
Touchon JC, McCoy MW (2016) The mismatch between current statistical practice and doctoral training in ecology. Ecosphere 7(8):e01394
Underwood AJ (1997) Experiments in ecology: their logical design and interpretation using analysis of variance. Cambridge University Press, Cambridge
Van Der Maaten L, Postma E, Van den Herik J (2007) Dimensionality reduction: a comparative review. J Mach Learn Res 10(66–71):13
Warren VE, Marques TA, Harris D, Thomas L, Tyack PL, Aguilar de Soto N, Hickmott LS, Johnson MP (2017) Spatio-temporal variation in click production rates of beaked whales: implications for passive acoustic density estimation. J Acoust Soc Am 141(3): 1962–1974
Wasserstein RL, Schirm AL, Lazar NA (2019) Moving to a world beyond "p<0.05". Taylor & Francis, New York
Wilcox RR (2010) Fundamentals of modern statistical methods: substantially improving power and accuracy. Springer, New York
Williams PJ, Hooten MB (2016) Combining statistical inference and decisions in ecology. Ecol Appl 26: 1930–1942
Xie Y, Allaire JJ, Grolemund G (2020) R markdown: the definitive guide. CRC Press, Boca Raton, FL
Yoccoz NG (1991) Use, overuse, and misuse of significance tests in evolutionary biology and ecology. Bull Ecol Soc Am 72(2):106–111
Zimmer WM, Johnson MP, Madsen PT, Tyack PL (2005) Echolocation clicks of free-ranging Cuvier's beaked whales (*Ziphius cavirostris*). J Acoust Soc Am 117(6):3919–3927
Zuur A, Ieno EN, Smith GM (2007) Analyzing ecological data. Springer, New York
Zuur A, Ieno EN, Walker N, Saveliev AA, Smith GM (2009) Mixed effects models and extensions in ecology with R. Springer, New York

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



# Behavioral and Physiological Audiometric Methods for Animals

10

Sandra L. McFadden, Andrea Megela Simmons, Christine Erbe, and Jeanette A. Thomas

## 10.1 Introduction

Audiometric studies, using behavioral or physiological methods, describe and quantify the hearing capabilities of animals. Audiometric studies using behavioral methods test hearing directly, by requiring an animal to make an observable response when it hears a target sound. The required response can be a natural, untrained response to sound, or the response can be one the animal is trained to make using classical or operant conditioning procedures. Physiological audiometric data, which do not require training, are more easily obtained than are behavioral data based on conditioning procedures. However, physiological methods can assess the perceptual process of hearing only indirectly. If it is shown that an animal's auditory system is capable of responding to sounds, the ability to hear may be inferred but is not guaranteed. For this reason, behavioral methods are considered the "gold standard" for audiometric assessment.

Animals hear sounds across a range of frequencies, and their sensitivity to audible sounds varies with frequency. By employing behavioral or physiological methods, researchers can determine the range of sound frequencies that animals hear, the amount of energy needed for the detection of sounds at each frequency, and the particular sound frequencies to which animals are most sensitive. Determining what sounds animals hear provides information about their acoustic environment and insight into the evolution of hearing among taxa. For example, toothed whales, microchiropteran bats, some shrews, and oil birds have evolved hearing abilities adapted for echolocation (see Chap. 12 on echolocation and the taxon-specific chapters in upcoming Volume 2), and some insect and fish prey have evolved keen hearing to detect their echolocating predators. Sounds to which animals are most sensitive are the ones most relevant to intraspecies communication and survival (because they provide information about mating partners or about predators and other sources of danger) and therefore are of particular interest.

In addition to providing information about normal hearing capabilities of animals, audiometric studies can show how hearing changes as a function of aging, environmental challenges, and experimental manipulations. Like humans,

Jeanette A. Thomas (deceased) contributed to this chapter while at the Department of Biological Sciences, Western Illinois University-Quad Cities, Moline, IL, USA

S. L. McFadden (✉)
Department of Psychology, Western Illinois University, Macomb, IL, USA
e-mail: sl-mcfadden@wiu.edu

A. M. Simmons
Department of Cognitive, Linguistic, & Psychological Sciences, Brown University, Providence, RI, USA
e-mail: andrea_simmons@brown.edu

C. Erbe
Centre for Marine Science and Technology, Curtin University, Bentley, WA, Australia
e-mail: c.erbe@curtin.edu.au

© The Author(s) 2022

C. Erbe, J. A. Thomas (eds.), *Exploring Animal Behavior Through Sound: Volume 1*,
https://doi.org/10.1007/978-3-030-97540-1_10

animals can experience presbycusis (i.e., loss of hearing with age; Willott 1991; McFadden et al. 1997) and they can develop hearing loss if exposed to ototoxic drugs, such as aminoglycoside antibiotics or platinum-based anti-cancer medications (Henderson et al. 1999). Hearing loss in wildlife due to noise exposure is of increasing concern because of widespread noise sources associated with anthropogenic activities in the ocean and on land (see Chap. 13 on the effects of noise). Audiometric studies of animals can also contribute to the understanding and treatment of human hearing and hearing disorders. For example, the study of the genetic and biological bases of hearing disorders often involves audiometric testing of animals with induced genetic conditions (e.g., knockin and knockout mice in which an existing gene is replaced or disrupted with an artificial piece of DNA, thereby altering or eliminating its function) and the investigation of pharmacological influences on human hearing is studied in laboratory animals.

Audiometric studies have been conducted on many aquatic and terrestrial species, with the choice of species guided by availability and the particular questions (biological, medical, or evolutionary) that the experimenter poses. Hearing abilities have been studied extensively in traditional laboratory mammals (Fig. 10.1) including the house mouse (*Mus musculus*), chinchilla (*Chinchilla lanigera*), Mongolian gerbil (*Meriones unguiculatus*), guinea pig (*Cavia porcellus*), and laboratory rat (*Rattus norvegicus*). These species are easy to obtain, easily bred in the laboratory, and readily trained in conditioning procedures, and so have long served as models for both normal and impaired human hearing. Audiometric studies have been conducted with many non-mammal species, including insects, amphibians, reptiles, fishes, and birds (see Volume 2). Many species are challenging to obtain, to house, and to train in a laboratory environment. For these reasons, behavioral audiograms are sometimes based on data from only one or very few animals, which limits the generalizability of the results. Further, hearing in some species is estimated by phonotaxis and evoked calling methods, which do not require training but which likely underestimate the animals' true hearing sensitivity. Understanding the auditory capabilities of non-traditional species provides insight into how hearing has become adapted to the challenges that animals face in a variety of natural environments. Unfortunately, for the vast majority of species, and even major taxa, there are no audiometric data available.

## 10.2 What Is an Audiogram?

An audiogram is a graph of hearing threshold as a function of frequency (ANSI/ASA S3.20-2015; ISO 18405: 2017).[1] Frequency refers to the sinusoidal vibration in cycles/s of a pure tone (sine wave). The hearing threshold of a listener is defined as the minimum stimulus level that evokes an auditory sensation in a specified fraction of trials at a given frequency. On an audiogram (Fig. 10.1), low threshold values correspond to high sensitivity to sound at that frequency and vice versa. The stimulus level is often a root-mean-square sound pressure level (SPL) expressed in dB with a reference of 20 μPa when testing in air or 1 μPa when testing under water; see Chap. 4, Introduction to Acoustics. The stimulus level may also be a root-mean-square sound particle velocity level (e.g., in the case of some fish audiograms) specified in dB re 1 nm/s. Because audiograms may be measured with signals other than pure tones (e.g., tone pips or clicks), signal type, threshold level, and reference value should be reported, along with the measured ambient noise levels. If the ambient noise is negligible, the hearing threshold is referred to as an unmasked threshold. If the ambient noise is high enough to raise the hearing threshold above its unmasked level, the hearing threshold is called a masked threshold (ISO 18405: 2017).

[1] Acoustical Society of America, Standard Acoustical & Bioacoustical Terminology Database: https://asastandards.org/asa-standard-term-database/; accessed 5 January 2021.

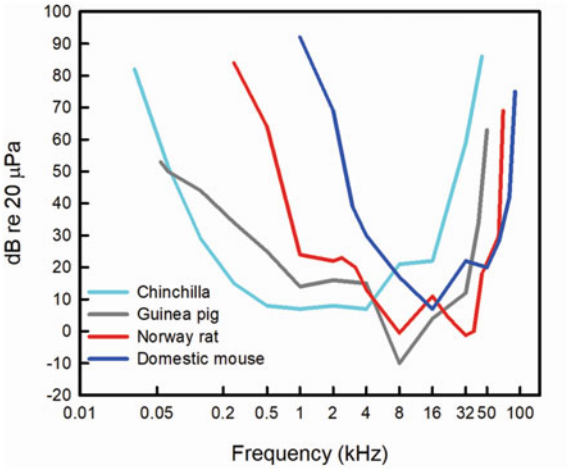





**Fig. 10.1** Left: Behavioral audiograms of rodents commonly used as laboratory animal models for hearing. Tones were presented through loudspeakers, and the animals' conditioned responses measured. All of the audiograms are U-shaped, with frequencies of best sensitivity (tip of the audiogram, at the lowest sound pressure level) within the range of 4–16 kHz. These species differ considerably in the low-frequency limit of hearing, with the chinchilla being more sensitive to a broader range of low frequencies than the domestic mouse. Plots are averaged thresholds based on 50% correct detection. Data were collected by Heffner and Heffner (1991, from three chinchillas); Koay et al. (2002, from two domestic mice); Heffner et al. (1994, from four Norway rats); and Heffner et al. (1971, from four Mongolian gerbils). Right: The photo of a mouse participating in a behavioral hearing test is courtesy of Micheal Dent, University at Buffalo, The State University of New York (Screven and Dent 2019)

There are two general approaches to assessing the auditory thresholds of live animals: behavioral and physiological. The behavioral hearing threshold is the lowest level that evokes a behaviorally measurable auditory sensation in a specified fraction of trials (ISO 18405: 2017). The pure-tone behavioral hearing threshold measurement procedure (prescribed in ANSI/ASA S3.21-2004) recommends that the behavioral hearing threshold be defined as the lowest input level at which responses occur in at least 50% of a series of ascending trials (i.e., trials in which signal level is systematically increased). The behavioral hearing threshold provides an integrated, whole-organism response to signal detection.

An electrophysiological hearing threshold is the lowest level that evokes a detectable and reproducible electrophysiological response (ISO 18405:2017). Both the ambient noise and the background electrophysiological noise levels should be reported. Electrophysiological noise is the non-acoustic self-noise arising from myogenic and neurogenic sources plus any artifact due to non-biological electrical interference. Electrophysiological hearing threshold estimates can be determined from different physiological processes (e.g., microphonic potentials, auditory brainstem response, cortical evoked responses), which characterize auditory processing at different levels of the auditory system. Various threshold estimation procedures also exist; each carries with it associated errors and assumptions, so the method for threshold estimation should be specified.

Electrophysiological methods are not equivalent to behavioral procedures, and electrophysiological hearing thresholds can differ from behavioral hearing thresholds (even for the same test animal). Within each of these two approaches, several methods can be employed, depending on the species being tested and the goals of the researcher. Behavioral techniques can be based on either unconditioned responses

that the animal makes spontaneously and as part of its natural repertoire, or conditioned responses that the animal is trained to make. Common physiological techniques measure otoacoustic emissions (OAEs; i.e., sounds generated by outer hair cells in the inner ear and measured using a very sensitive microphone) and auditory evoked potentials (AEPs; i.e., summed electrical responses of hair cells and auditory neurons recorded from electrodes). Results from behavioral and AEP experiments in the same species or even in the same animal can produce audiograms that are similar in shape and frequency range but may differ in absolute thresholds (see Sect. 10.4.3).

Audiograms in most species are typically U-shaped, but not symmetrical (Fig. 10.1). The frequency region of best sensitivity encompasses those sound frequencies at the trough of the U-shaped curve, where thresholds are lowest. The animal's best hearing sensitivity (or lowest threshold) corresponds to the threshold range at the frequency region of best sensitivity. The range of hearing specifies the sound frequencies that are audible to an animal at some specified level (e.g., 60 dB) above the lowest threshold. The range of hearing for sounds at high sound levels is wider than the range of hearing for sounds at low sound levels because the audiogram is broad and U-shaped. The range of hearing should be expressed as between X Hz and Y Hz at Z dB above the best hearing sensitivity. Unfortunately, many publications do not include the number of decibels above the best hearing sensitivity when reporting the range of hearing for an animal or species, and they may not indicate whether the highest and lowest frequencies shown in an audiogram reflect the limits of testing or the limits of the animal's hearing capabilities.

In terrestrial mammals, the main contributors to the U-shape of the audiogram and the location of the frequency of best sensitivity are the acoustic properties of the auditory periphery: the pinnae, external auditory meatus, and middle ear (Tonndorf 1976; Hellström 1995). The pinna serves to funnel sounds into the external auditory meatus (i.e., the ear canal), with sounds from some directions being amplified and those from other directions being attenuated. The external auditory meatus is an acoustic resonator that boosts the amplitude of received frequencies at and near its resonant frequency. The resonant frequency of the ear canal is inversely proportional to its length, so animals with short ear canals, such as mice, have their best hearing sensitivity at high frequencies, whereas animals with long ear canals, such as elephants, have their best hearing sensitivity at low frequencies. The resonant characteristics of the external auditory meatus, coupled with the sound transfer properties of the middle ear, help determine the acoustic energy levels reaching the inner ear.

Often, audiograms are incorrectly interpreted as illustrating hard thresholds to sounds, assuming that sounds at amplitudes just below the published audiogram are inaudible and sounds just above the audiogram are always audible. That is not the case. The faintest sound that an animal can hear depends on many factors, including stimulus characteristics (e.g., duration, repetition rate), environmental factors (e.g., ambient noise level, testing context such as anechoic chamber versus natural environment), and individual factors (e.g., health, response bias, attention, age). A given animal may show a loss of sensitivity due to aging, noise exposure, or exposure to ototoxic drugs, and even due to repeated or prolonged exposure to the stimulus during testing that leads to sensory adaptation and/or cognitive habituation. At high ambient noise levels or when additional sounds are present, an animal might lose the ability to hear a sound it previously heard in a quiet environment. This is because of masking, in which the presence of non-target sounds or noise decreases the detectability of the sound of interest.

Within a species, there can be significant individual differences in hearing sensitivity, which can reflect differences in attention to the task, age, health, and history of exposure to sounds, among other factors. Because there can be considerable variability among animals of a given species, it is important to test many animals when possible. Also, it is important to know when examining an audiogram whether the

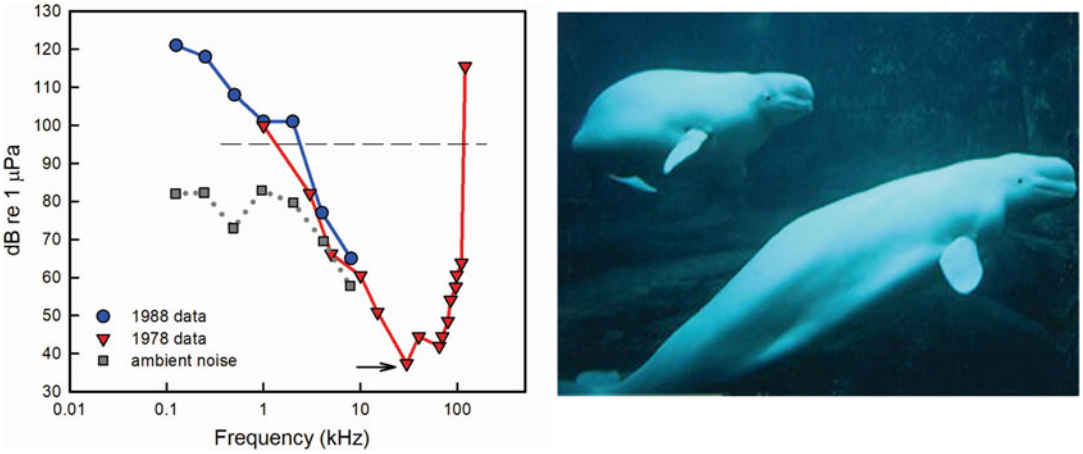



**Fig. 10.2** Left: Underwater behavioral audiograms of three beluga whales obtained at two different times 10 years apart. Data were obtained using an ascending Method of Limits (described in Sect. 10.3.3). The whales were trained to leave a station when they heard a tone and swim to the trainer for a food reward. Thresholds were defined as the tone level at which the whales detected the signal 50% of the time. The red triangles show the mean audiogram from one male and one female beluga whale reported by White et al. (1978). The arrow shows the most sensitive frequency at 30 kHz. The blue circles show averaged data from the same male and female and an additional juvenile male, obtained by Awbrey et al. (1988). The gray squares show the ambient noise level in the test pool, which was close to the measured thresholds at 4 and 8 kHz, indicating that the whales' actual thresholds at these frequencies were likely lower than indicated on this graph. The gray dashed line is 60 dB above the lowest threshold at 30 kHz, where the range of hearing was measured. Right: Photo of two beluga whales at Vancouver Aquarium

curve is based on a single animal or a group of animals.

Audiograms from three beluga whales (*Delphinapterus leucas*) are shown in Fig. 10.2. From this graph, it can be seen that testing was conducted in water because the dB reference is 1 µPa, rather than 20 µPa for sounds presented in air (as in Fig. 10.1). In belugas, hearing sensitivity increased from low frequencies around 250 Hz to the best frequency range around 30 kHz (threshold around 37 dB re 1 µPa), and then decreased toward higher frequencies up to 120 kHz; this results in a U-shaped hearing curve. The range of hearing at 60 dB above lowest threshold extends from about 1–110 kHz.

## 10.3 Behavioral Methods for Audiometric Studies on Live Animals

Behavioral approaches can be divided into two general types, unconditioned response techniques and conditioned response techniques. Unconditioned response techniques are based on behaviors that the animal naturally makes to sound and are readily employed in the animal's natural habitat. Animals must be trained to make conditioned responses, and this training should be based on the species' typical behavioral repertoire. Klump et al. (1995) provide a full discussion of different methods used to study hearing sensitivity in animals.

For both techniques, establishing stimulus control over an animal's behavior is crucial. A pure tone is typically the test signal, although broadband clicks, and noises of varying bandwidths can be used, depending on the research question. How signals are generated and presented is extremely important to control and monitor. The sound may be delivered via a loudspeaker to animals ranging freely, being confined to the experimental chamber, or trained to hold station (e.g., at a bite plate or in a hoop), or delivered via tubes, insert earphones, or headphones (Fig. 10.3). Stimuli can be presented

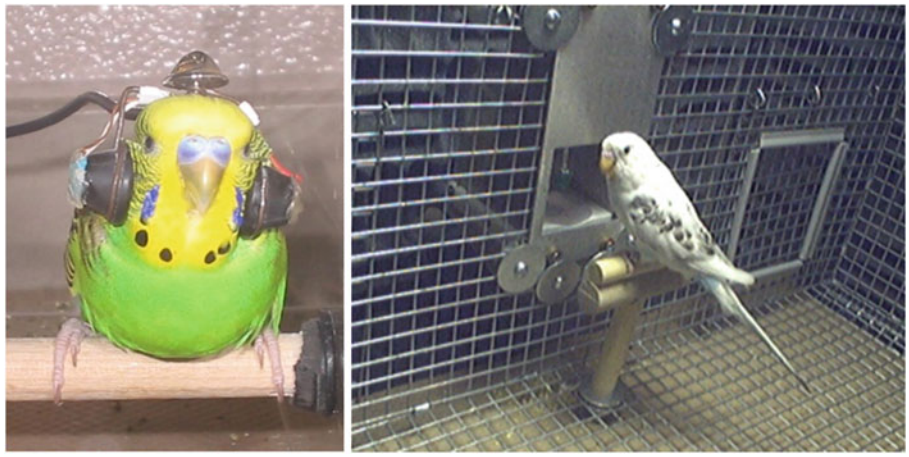


**Fig. 10.3** Photos of a budgerigar (*Melopsittacus undulatus*) wearing headphones during a sound localization experiment (left; Welch and Dent 2011) and receiving a reward during a frequency discrimination experiment (right; Dent et al. 2000). Courtesy of Micheal Dent, University at Buffalo, The State University of New York

using several different protocols, each of which has its own assumptions and limitations. Ambient noise can influence thresholds and so must also be controlled. Ambient noise can be minimized if the animal is tested in an anechoic chamber or a sound-attenuating chamber (Fig. 10.4). If animals are tested in their natural environments where ambient noise levels cannot be controlled, researchers must take periodic measurements of the amount of ambient noise present during hearing tests.

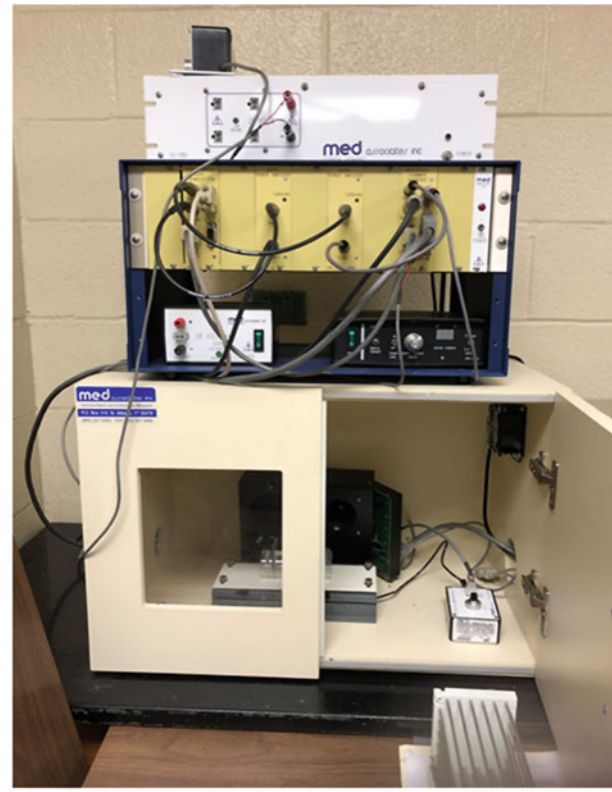
### 10.3.1 Behavioral Methods Using Unconditioned Behaviors

#### 10.3.1.1 Preyer Reflex and Acoustic Startle Response

The Preyer reflex and the acoustic startle response (ASR) are behaviors triggered automatically by unexpected, high-amplitude sounds. These are reflexive responses to sound that require no training of the animal and thus are relatively easy to implement. On the other hand, animals can habituate to repeated presentations of high-amplitude sounds that best evoke these reflexes. Thus, sound-evoked reflexes can be useful as fast and easy screening tests for bracketing an animal's hearing abilities but are not good measures for determining absolute thresholds of hearing.

The Preyer reflex has been described as an orientation or attentional reflex (Jero et al. 2001). In mammalian species that are able to move their pinnae, it involves a quick retraction of the ears, a rapid twitch of the ears, or a change in orientation of the pinnae toward the source of the sound. In species with immobile pinnae, turning of the head toward the sound source (which brings the source of the sound into the animal's line of vision) is the measure of orientation. In some studies, a trained observer simply rates the Preyer reflex as present or absent. The reflex also can be monitored using a motion-tracking camera system and reflective markers attached to each of the animal's pinnae, as described in a study using the guinea pig (Berger et al. 2013). The magnitude and latency of the Preyer reflex can then be determined by measuring pinnae displacement during sound presentation.

The ASR is a whole-body response to unexpected sounds presented at very high amplitudes (typically above 90 dB re 20 μPa) and has been interpreted as a protective or alarm reflex. It can be elicited in a wide range of adults and developing vertebrates, including fishes and most mammals, and typically is quantified in terms of



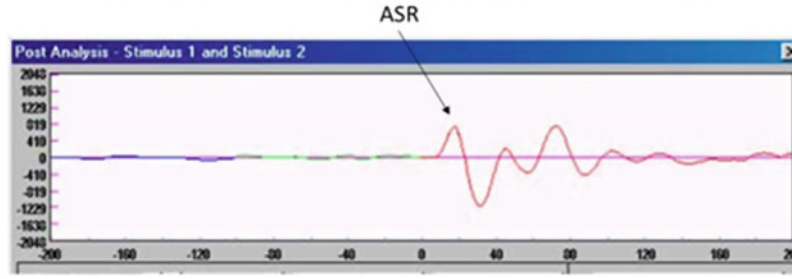




**Fig. 10.4** A sound attenuating chamber set up for acoustic startle reflex (ASR) testing in small animals such as mice and rats. The animal is placed in a plastic tube or a wire restraining device on an accelerometer platform. Voltages produced by the movement of the animal on the platform are recorded and quantified. Typical ASR measures are peak amplitude and response latency

response amplitude and response latency. In teleost fish, the ASR is called the tail-flip reflex or C-start response, and it involves an initial full flexion of the body followed by a weaker flexion in the opposite direction, so that the animal bends and swims away from the source of the stimulus. The response is mediated by the Mauthner cells, a pair of giant neurons located at the level of the auditory-vestibular nerve in the hindbrain. The Mauthner cells receive input from the auditory nerve and then send signals to motor neurons on the opposite side of the body, which then produce the behavioral response. The ASR in fishes can be measured by placing the animals in small acrylic plates filled with water and mounted on top of a vibration device that produces particle motion stimulation. A high-speed video camera is needed to visualize the C-start response (Bhandiwad and Sisneros 2016).

In small mammals such as rodents, the ASR consists of hunching of the shoulders, dorsiflexion of the neck, and rapid extension then flexion of the limbs. ASR in rodents is typically measured by placing the animal on a platform that measures displacement and force or acceleration caused by limb extension (Fig. 10.4). In primates, the ASR involves the reflex contraction of striate skeletal muscles,

primarily muscles of the face, neck, shoulders, and arms (Braff et al. 2001).

An animal that twitches its ears or startles repeatedly (e.g., in at least two out of three presentations) in response to finger snaps, hand claps or pure tones at different frequencies has demonstrated an ability to hear. At the same time, however, the presence of a startle response does not mean the animal has normal hearing. This was demonstrated clearly in a study of the sensitivity and specificity of the Preyer reflex by Jero et al. (2001). The researchers used hand claps or the metallic sound of two hammers hitting together to elicit startle responses from young adult albino laboratory mice of the FVB strain. They found that the reflex test was effective for identifying profound hearing loss, but was insensitive for identifying less severe hearing losses.

Reflex responses to sound can be used to show differences between groups of animals as a function of age or experimental treatment. Bhandiwad and Sisneros (2016) examined the development of hearing in two species of larval fishes, the three-spined stickleback (*Gasterosteus aculeatus*) and the zebrafish (*Danio rerio*), by quantifying the probability of a startle reflex in response to sounds of different frequencies at different ages post-fertilization. McFadden et al. (2010) showed declines in the amplitude and increases in the latency of the ASR with age in laboratory rats. Age-related changes in one or more of the components of the ASR circuit or to brain regions providing inhibitory input to this circuit can account for ASR changes observed in older animals and humans.

Startle responses also can be useful for determining the range of frequencies that an animal can hear. Bowles and Francine (1993) determined that kit foxes *(Vulpes macrotis*) have a functional hearing range from 1 to 20 kHz by observing startle responses of four wild-caught kit foxes to playbacks of tones of different frequencies. An additional advantage of startle reflex testing is that a group of animals can be tested simultaneously. Kastelein et al. (2008) determined the frequency range of hearing for eight species of marine fish by noting the frequencies at which 50% or more of the fish in a school reacted to the sound stimulus by increasing swimming speed and making tight turns. Disadvantages of using startle responses are that they require presentation of high amplitude stimuli and they habituate quickly.

#### 10.3.1.2 Prepulse Inhibition (PPI) and Reflex Modification

Although the ASR is a reflex that is not typically under voluntary control, it is sensitive to and can be modified by ongoing behaviors and attentional status of an animal. The ASR can be potentiated under some circumstances and attenuated or inhibited under others. Animals typically show larger ASRs when they are afraid or anxious than when they are not, so fear-potentiated startle paradigms commonly are used to study fear and anxiety states in animals. When an animal is processing another stimulus, such as a brief low-level sound or a puff of air or a flash of light, it will startle less to a sudden, loud sound than when it is not otherwise engaged. The ability of an auditory, tactile, or visual prepulse stimulus to reduce the amplitude of the ASR is termed prepulse inhibition (PPI).

Even an auditory prepulse stimulus near the hearing threshold of an animal can attenuate the ASR, and this makes the PPI paradigm suitable for testing threshold levels of sound and determining subtle effects of treatments on auditory function. PPI has been used to study the auditory sensitivity of fishes, frogs, and mammals (Fig. 10.5). In larval zebrafish, the probability of an ASR to a high-amplitude tone was reduced when the tone was preceded by other tones at sub-startle levels (Bhandiwad and Sisneros 2016). Thresholds obtained by PPI in this species were lower than thresholds obtained by using the ASR alone.

Reflexes other than acoustic startle responses can be modified by the prior presentation of a sound; these paradigms are termed reflex modifications (Hoffman and Ison 1980). Simmons and Moss (1995) adapted this paradigm to obtain audiograms for two species of frogs, the American bullfrog (*Lithobates catesbeianus*) and

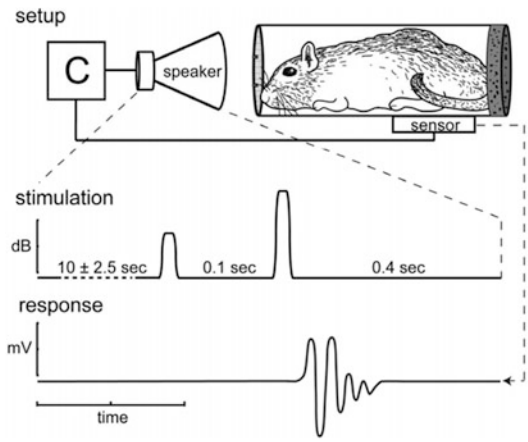



**Fig. 10.5** Schematic drawing of a setup used to study prepulse inhibition of the ASR in Mongolian gerbils. The top drawing shows a gerbil placed into an acrylic tube 10 cm in front of a loudspeaker. The force sensor under the acrylic tube monitors the gerbil's movements. The C label shows the position of the stimulation/recording computer. Center drawing shows the timing of acoustic stimulation (dB) with the pre-stimulus (lower amplitude trace) preceding the startle-producing stimulus (higher amplitude trace). Bottom drawing shows the response measured by the force sensor. Here, the response occurs only to the stimulus and not to the pre-stimulus. After repeated pairings of the pre-stimulus and stimulus, the response to the stimulus declines (Walter et al. 2012). © Walter et al. 2012; https://www.scirp.org/journal/paperinformation.aspx?paperid=17796. Licensed under CC BY 4.0; https://creativecommons.org/licenses/by/4.0/

the green treefrog (*Dryophytes cinereus*). Frogs were constrained inside a small dish (1–2 cm in diameter larger than the animal), which was then placed on top of a stabilimeter that picked up the frog's movements within the dish. Two copper strips cemented to the side of the dish produced a mild electric shock that evoked small reflex contractions of the frog's hind limbs. The reflex evoked by the electric shock was modified in strength by prepulses of pure tones, with the extent of modification varying with prepulse amplitude. At any given tone frequency, the amplitude of the prepulse producing 10% inhibition of the reflex response was defined as the threshold to that frequency. The magnitude of the reflex modification effect varied with the amplitude of the prepulse, but only when stimulation was spaced at intervals wide enough to avoid habituation.

#### 10.3.1.3 Phonotaxis

Some animals have a natural tendency to approach sound (positive phonotaxis) or make evasive movements away from sound (negative phonotaxis). Sounds that elicit positive phonotaxis include species advertisement calls (i.e., mating calls), while sounds that elicit negative phonotaxis include sounds made by predators. These natural behavioral responses to sound can be exploited to estimate hearing sensitivity in those species for which training procedures based on conditioned responses are extremely difficult to implement. Phonotaxis experiments are readily conducted in the animal's habitat and so can provide crucial information on the acoustic features animals use to recognize conspecific (own species) vocal signals such as advertisement and aggressive calls. These kinds of field studies are particularly important for identifying the impact of the entire soundscape on sound detection and discrimination, and for assessing the effects of environmental variables, such as air temperature and humidity, on acoustic communication.

Phonotaxis has been especially useful for studying auditory capabilities of female orthopteran insects, frogs, and songbirds, because these animals naturally approach stationary calling males in order to mate with them. For example, gravid female frogs readily approach loudspeakers broadcasting sounds (tone bursts, amplitude-modulated tones, or frequency-modulated tones) which they recognize as components of the advertisement calls of males of their own species, or even a synthetic version of these conspecific calls (Gerhardt 1995). The sensitivity of females to these sounds is measured in experiments in which sounds of different levels, frequencies, or temporal patterning are broadcast from a loudspeaker, and the female's approach to the loudspeaker is quantified. Sounds can be broadcast from one source (one-speaker design) to estimate sound detection or from two sources (choice or two-speaker design) to estimate sound discrimination. The researcher can obtain an estimate of the female's relative sensitivity to sounds (if sound frequency is varied) or her ability to distinguish sounds of two intensities

(if sound level is varied). Responses are quantified in terms of the nearness and the path of the phonotactic approach, the latency of the response, and the presence of orientation movements, such as head-turning toward the sound source. Data are typically presented as the proportion of females responding to a particular stimulus as a function of whatever parameter is being varied, with the 50% correct point on the resulting function defined as the threshold in a one-choice experiment and the 75% correct point (midway between chance and perfect performance) defined as the threshold in a two-choice experiment (see Volume 2, Chap. 3 on amphibians).

Because most species of insects and frogs call at night, visualizing their movements in a phonotaxis experiment can be challenging. Figure 10.6 shows a new technique designed to monitor phonotactic movements of frogs in both the laboratory and the natural environment (Aihara et al. 2017). In this technique, a female Australian orange-eyed treefrog (*Ranoidea chloris*) wears a miniature LED backpack. A video camera records the energy emitted from the LEDs, thus allowing researchers to track the frog's movements. Sounds are broadcast through multiple loudspeakers, and monitored by separate LED sound indication devices, each of which has a different pattern of illumination. In this way,

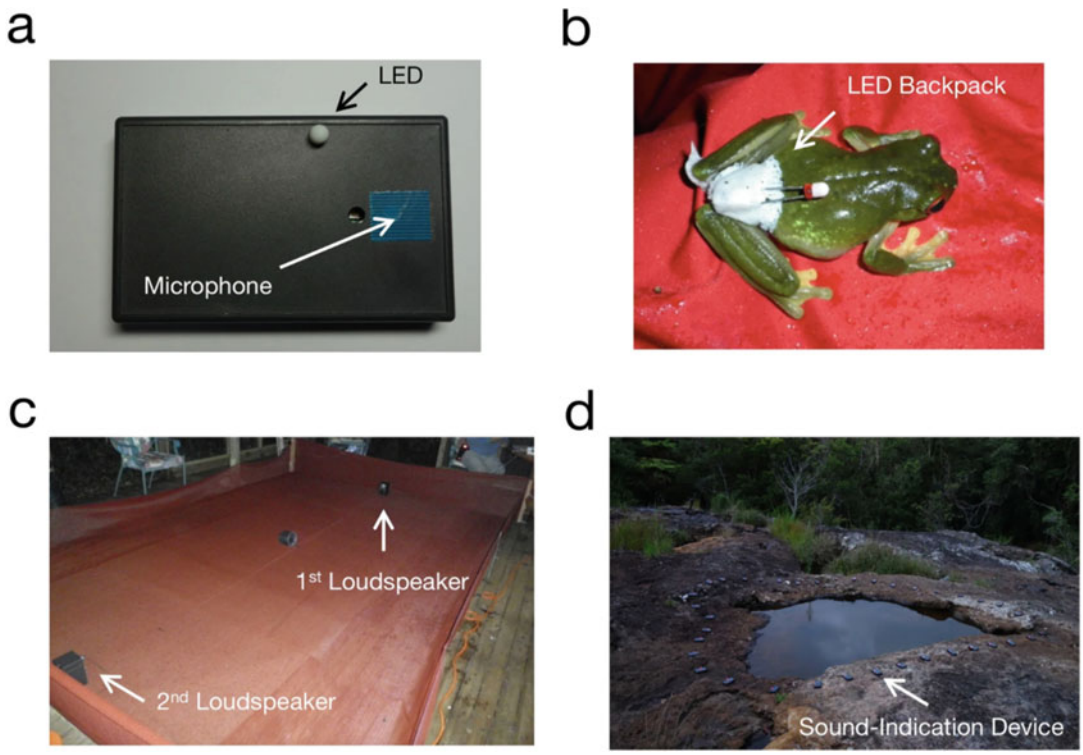



**Fig. 10.6** (**a**) An image of a sound indication device that consists of a miniature microphone and a light-emitting diode (LED). The LED is illuminated when detecting sounds. (**b**) Photo of an orange-eyed female treefrog wearing a LED backpack. (**c**) Arena playback experiment. Two loudspeakers at each end of the arena present sounds. A sound indication device is placed in front of each loudspeaker. The female wearing the backpack is released from the middle of the arena. The lights emitted by the sound indication device and the LED backpack are recorded by a video camera. (**d**) Natural habitat of the orange-eyed treefrog. The position of the sound-indication device is shown (Aihara et al. 2017). © Aihara et al. 2017; https://www.nature.com/articles/s41598-017-11150-y. Licensed under CC BY 4.0; https://creativecommons.org/licenses/by/4.0/

researchers can not only track the female's movements but also which of several loudspeakers is playing the preferred sound.

There are limitations to the use and interpretation of phonotaxis data. Although phonotaxis experiments can tell us which sounds animals prefer and how sensitive they are to these sounds, they are not suitable for the compilation of entire audiograms or estimates of an animal's entire range of hearing. When a female fails to approach a sound source, it may be because she does not hear it or because she does not recognize it as an advertisement call. Moreover, females of many species will show phonotaxis only when they are gravid. This limits the timespan during which experiments can be conducted, although phonotaxis can be induced by hormone injections (Gerhardt 1995). Male insects and frogs typically exhibit phonotaxis only in response to a high amplitude sound resembling an advertisement call or an aggressive call from a rival male. Males treat aggressive calls from rivals as threats and respond aggressively, by approaching the source and attempting to engage it physically. Because males are less likely than females to approach sound sources, descriptions of their hearing sensitivity based on phonotaxis are not reliable.

#### 10.3.1.4 Evoked Calling

Evoked calling is another method based on unconditioned responses that can be used to estimate hearing sensitivity and acoustic preferences. Males of some species (orthopteran insects, frogs, songbirds) vocalize in response to playbacks of signals resembling conspecific advertisement or aggressive calls. The male's sensitivity to these playbacks can be estimated by lowering the amplitude of the signal until the male no longer vocalizes back. Varying the acoustic features (frequency, temporal patterning) of the signal can provide estimates of sensitivity to these particular features (Fay and Simmons 1999). Evoked calling experiments, like phonotaxis experiments, can be implemented either in the laboratory or in the field. As with the phonotaxis technique, the evoked calling technique does not measure audibility per se but can be useful for determining what acoustic features of communication signals are most important for mediating behavioral responses. Despite their limitations, phonotaxis and evoked calling techniques are useful because they provide insight into what sounds animals pay attention to in their natural environment and thus into perceptual decision-making in a biologically relevant context.

### 10.3.2 Behavioral Methods Using Conditioned Behaviors

#### 10.3.2.1 Classical Conditioning

Classical conditioning techniques have been used to train several species of animals for audiometric studies. In classical conditioning, an unconditioned stimulus that naturally elicits an unconditioned response is paired with a conditioned stimulus. After a number of pairings of the conditioned stimulus with the unconditioned stimulus, presentation of the conditioned stimulus alone elicits a conditioned response that is the same as or similar to the unconditioned response.

Fay (1995) described the use of classical respiratory conditioning to estimate auditory thresholds in the goldfish (*Carassius auratus*). The goldfish was restrained in a cloth bag and submerged in a small tank. An underwater loudspeaker was placed on the bottom of the tank. A tone of a particular frequency was presented shortly before a brief electric shock (unconditioned stimulus) that produced an unconditioned suppression of the fish's respiration. Changes in the amplitude and rate of fish's respiration were measured by a thermister placed in front of the fish's mouth. After multiple pairings of the tone and shock, presentation of the tone alone produced a conditioned suppression of respiration. By determining the amplitude level of the tone that no longer produced a conditioned response, the fish's sensitivity to that tone frequency could be determined.

Ehret and Romand (1981) used both unconditioned and classically conditioned pinnae movements and eye-blink responses to track the postnatal development of auditory thresholds in domestic kittens (*Felis catus*). Unconditioned

movements of the pinnae and/or facial muscles in response to high-intensity tone bursts were observed in one group of kittens up to 12 days of age. A second group of kittens (aged 10 days to 1 month) was trained with tone-shock pairs to make conditioned movements of their eyelids and pinnae when they heard a sound. Ehret and Romand's results showed that some kittens as young as 1–2 days of age were able to respond to some frequencies, and that sensitivity to low, mid, and high frequencies developed at different ages.

#### 10.3.2.2 Operant Conditioning

There are many responses animals can make to indicate when sounds are heard (or not heard), such as touching a response paddle, pressing a lever with a nose or paw, lifting a paw, licking a tube from a water bottle, swimming across a barrier, or vocalizing. It is important to choose a response that is based on an animal's natural behaviors and thus is easy to learn. Once the response is chosen, there are several behavioral methods that can be used to train animals to make the response when a sound is detected or refrain from the response when no stimulus is presented. These different paradigms have been implemented successfully with a large number of species, with modifications that take into account species-typical behaviors and habitats.

Operant conditioning techniques can use positive or negative reinforcement procedures for training or "shaping" a conditioned response. Positive reinforcement methods establish the behavior by providing a reward, such as food, water, or even verbal praise or tactile stimulation whenever the animal makes the appropriate response. Negative reinforcement methods remove an unpleasant or aversive stimulus (usually mild electric shock) whenever the animal makes the appropriate response. Methods can also be used to decrease unwanted or incorrect responses; these are termed punishment procedures. For example, a time-out period might be imposed (positive punishment) when an animal makes an incorrect response. After the desired behavior has been established through an appropriate schedule of reinforcement during a training phase, the animal is then tested using various frequencies and amplitudes of sound to determine the audiogram. Sometimes animals mistakenly respond when there is no signal present; this is a false alarm. Some animals are more inclined to make false alarms than others. To assess this bias, "catch trials" (i.e., control trials in which no signal is presented) are interspersed at random in the stimulus series. Some researchers desire to assess the animal's attentiveness to a hearing task before collecting data, such as by conducting a set of easily heard "warm-up trials" at the beginning of a session, and a set of easily heard "cool-down trials" at the end of a session. Criteria can be set such that if the animal's performance does not reach a certain percent of correct responses during either the warm-up or the cool-down trials (e.g., 80%), testing is discontinued for that session or data from that session are eliminated.

In conditioned suppression/avoidance paradigms, an animal learns to suppress an ongoing behavior when it detects a sound that signals shock (Heffner and Heffner 2001). The shock levels used in these studies are kept low so that the animals do not become agitated or develop a fear of the test apparatus that would impair their performance. Heffner et al. (2014) used the conditioned suppression procedure to determine behavioral audiograms and sound localization abilities of three young male alpacas (*Vicugna pacos*). Thirsty alpacas were trained to break contact with a water spout when they heard a tone or noise signal (a conditioned stimulus) that warned of impending shock (unconditioned stimulus) and to resume drinking water following a safety signal. The safety signal for tone threshold testing was a shock indicator light that turned off when shock was terminated. Hit rates (measuring the percentage of correct detections of sound, indicated by breaking contact with the water bowl when the tone signal was present) and false alarm rates (measuring the percentage of false alarms, indicated by breaking contact with the water bowl when no tone was present) were determined for each stimulus intensity. The pure-tone thresholds of the three alpacas showed little variability among individuals. Indeed, Heffner and Heffner (2001) argued that individual variation among animals is less when using

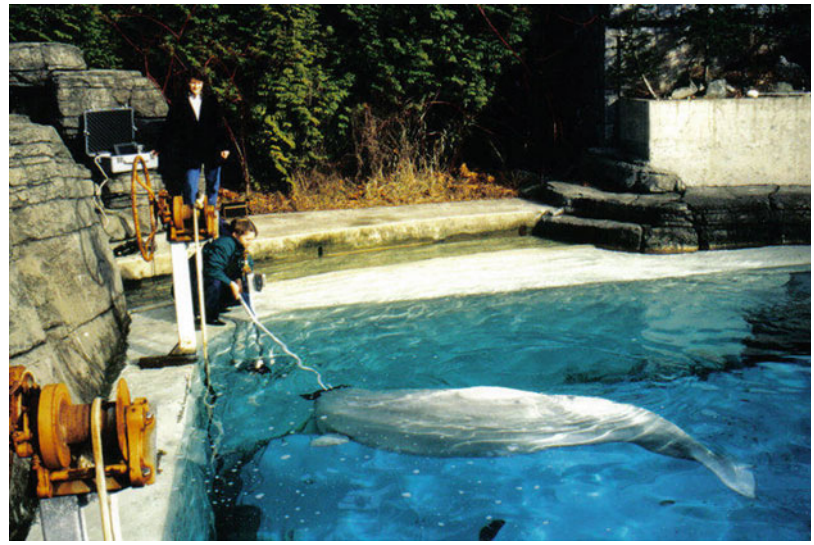


**Fig. 10.7** Photo of a beluga whale holding station in front of an underwater loudspeaker during behavioral training for later audiogram measurements at Vancouver Aquarium. During the actual experiment, the computer operator moved behind the rock wall, out of sight of trainers and whale

conditioned suppression compared to methods based on positive reinforcement.

Another common technique based on positive reinforcement, used in many species of aquatic (Fig. 10.7) and terrestrial species, is a go/no-go response paradigm. Thomas et al. (1990) used this technique to measure the audiogram of a subadult male Hawaiian monk seal (*Neomonachus schauinslandi*). At the start of each trial, a trainer sent the seal, using a hand cue, to station under water with its chin resting on a headstand. If a tone was heard, the seal was expected to leave the station, touch a response paddle, and swim to the trainer for a fish reward (go response). If no tone was heard (either a control trial or an inaudible signal), the seal was supposed to stay at the station, wait for the trainer to give a release whistle, and then swim back to the trainer for a reward (no-go response). Half the trials were signal-present and half were signal-absent controls; the order of presentation of the trial types was pseudorandomized throughout a session so that the animal would adopt a neutral response bias. The trainer then called the seal back to the initial station with a whistle and the next trial commenced.

There are several drawbacks of behavioral audiometric studies based on conditioning procedures. Most notably, weeks or months may be required to train the animal to respond reliably. It is important to maintain the animal's motivation to respond and attention to the task, both of which can wane if there are changes in the social environment, routine, or the animal's health.

Because behavioral audiograms require a long period to train and test the animal, and since the number of individuals in captivity is limited for many species, in some marine mammals, hearing data are available for only a single animal. Hall and Johnson (1972) conducted a behavioral audiogram on a captive killer whale (*Orcinus orca*) and reported that this species had much worse high-frequency hearing than other toothed whales tested to that date. Later, Bain et al. (1993) conducted behavioral audiograms on five killer whales and found their hearing was very typical of other toothed whales. Upon investigation, the researchers found that the original test subject had been given high dosages of an ototoxic antibiotic. So, the first killer whale tested was likely hearing impaired as a result of antibiotic-induced death of hair cells in the high-frequency region of the cochlea. By now, another eight individuals have been tested confirming more typical delphinid audiograms in killer whales (Branstetter et al. 2017).

### 10.3.3 Signal Presentation Paradigms for Behavioral Audiograms

There are three classic paradigms commonly used for signal presentation in behavioral audiogram

tests with animals (Levitt 1970; Klump et al. 1995): the Method of Constant Stimuli, the Method of Limits, and the Up/Down Staircase method (also called "adaptive tracking method"). One important factor to keep in mind when choosing a signal presentation paradigm is the time available for measuring thresholds, as there is a trade-off between the number of trials and the accuracy and reliability of hearing-threshold measurements.

#### 10.3.3.1 Method of Constant Stimuli

The Method of Constant Stimuli provides the greatest accuracy and reliability for threshold measurements. In this paradigm, the animal is tested at one frequency in a session with blocks of trials having an equal number of different signal levels ranging from very low to very high amplitude (i.e., no silent controls), presented in random order. The animal makes a response when a signal is heard, and the results for each signal presentation ("Yes" the tone was heard or "No" the tone was not heard) are tallied by amplitude levels (Fig. 10.8 left panel). After all responses are tallied, a psychometric function (i.e., a plot of the animal's responses, typically the percentage of "Yes" responses) versus amplitude level (Fig. 10.8 right panel) is made. The threshold level is determined (often by interpolation) as the level at which the animal indicated it heard the signal on 50% of the trials.

The stimulus presentation levels cover a wide range that bracket the animal's threshold, so additional points on the psychometric function can be estimated. Randomized presentation of stimuli prevents the animal from anticipating the stimulus level on the next trial. Many of the stimulus levels are well above threshold, so the animal is not required to make difficult detections on every trial. On the other hand, the method is time-consuming, and the choice of stimulus levels to present requires some prior knowledge of likely thresholds at a specific frequency.

#### 10.3.3.2 Method of Limits

The Method of Limits involves the presentation of stimuli in small steps (typically 2 to 5 dB) over a fixed range of stimulus levels. At each level, the experimenter records whether the animal responded to the test tone or not (Fig. 10.9). Stimuli may be presented in an ascending series, from the lowest amplitude to the highest, or in a descending series, from the highest amplitude to the lowest. Multiple runs are conducted, and for each run, the crossover level (i.e., the level halfway between the stimulus level not heard and the

| Stimulus Level | Number of Times Observer Said "Yes" (for 50 trials at each level) | Percent of total (for that level) |
|---|---|---|
| 1 | 3 | 6 |
| 2 | 4 | 8 |
| 3 | 5 | 10 |
| 4 | 11 | 22 |
| 5 | 21 | 42 |
| 6 | 38 | 76 |
| 7 | 48 | 96 |
| 8 | 50 | 100 |
| 9 | 50 | 100 |

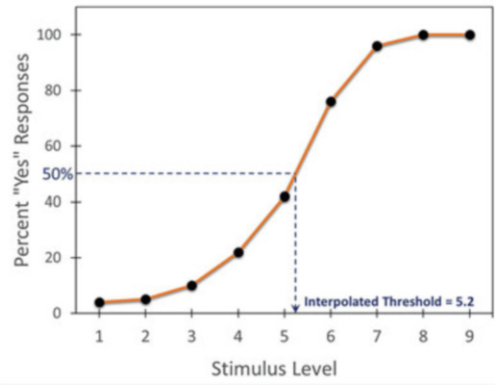



**Fig. 10.8** Illustration of the Method of Constant Stimuli. Left panel: Fifty stimuli were presented at each of nine stimulus levels (450 trials total). The number of times the subject indicated that the stimulus was heard at each level was tallied in the Number column and converted to a percentage in the Percent column. At stimulus levels below threshold, the subject rarely responded, whereas at the highest stimulus levels, the subject reported detection on all 50 trials (100%). Right panel: Data from the tallies chart were used to plot a psychometric function, showing performance as a function of stimulus level. Threshold, defined as the stimulus level at which the subject made a detection response on 50% of the trials, was interpolated to be 5.2 in this example

**Method Of Limits**

| Stimulus Level | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 40 | Y | Y | Y | Y | Y |
| 35 | Y | Y | Y | Y | Y |
| 30 | Y | Y | Y | Y | Y |
| 25 | Y | N | Y | N | Y |
| 20 | N | N | N | N | N |
| 15 | N | N | N | N | N |
| | 22.5 | 27.5 | 22.5 | 27.5 | 22.5 |

Threshold = average of crossovers = 24.5 dB re 20 µPa

Y = Yes response
N = No response

**Fig. 10.9** Illustration of the Method of Limits. Five series of trials (runs) were used, with test tones at six stimulus levels (15–45 dB re 20 µPa) presented in each run. Stimuli were presented from the highest level to the lowest (i.e., in descending order) on the first, third, and fifth runs, and from the lowest level to the highest (i.e., in ascending order) on the second and fourth runs. The crossover level was recorded for each run, then crossover levels were averaged to estimate threshold. In this example, a total of 30 trials were conducted across five runs, and the threshold was estimated to be 24.5 dB re 20 µPa

next level heard, e.g., 22.5 dB for run 1 and 27.5 dB for run 2 in Fig. 10.9) is determined. The mean threshold is estimated by averaging all of the crossover levels for that frequency.

Presenting all runs in either descending order or solely in ascending order may produce a strong response bias that influences threshold estimates. When trials are presented using the descending Method of Limits, the animal can become accustomed to reporting that it perceives a stimulus and can continue reporting hearing the signal below the threshold; this is known as the error of habituation. Alternatively, in the ascending Method of Limits, the animal can anticipate that the stimulus is about to become detectable and make an error in responding in the absence of the signal; this is known as the error of anticipation. The bias introduced by signal predictability is a drawback of using the Method of Limits. The influence of habituation and anticipation errors can be partly overcome by using an equal number of ascending and descending runs alternately on the same subject.

The Method of Limits is often preferred over the Method of Constant Stimuli because of its greater efficiency in bracketing thresholds; i.e., fewer trials are needed for a reliable estimate of threshold. In the example shown in Fig. 10.9, responses to test tones at six stimulus levels were recorded across five runs; this required 30 trials total. If the Method of Constant Stimuli had been used, with 50 signals presented at each of the six stimulus levels, a total of 300 trials would have been presented.

#### 10.3.3.3 Up/Down Staircase Method

The Up/Down Staircase method, or adaptive tracking signal presentation paradigm, is a variation of the Method of Limits that was developed by von Békésy (1960) as a way of efficiently determining thresholds (Fig. 10.10). This method is also referred to as a Modified Method of Limits. The test begins with the presentation of a high-amplitude signal that is likely to be easily heard. Then, the amplitude is reduced in 2- to 10-dB steps until the animal does not respond to the signal. When the animal signifies it can no longer hear the signal, the dB level is immediately increased (in 1- to 5-dB steps) until the animal reports it again hears the sound. At that level, the direction is reversed and the procedure is repeated. Thus, this method includes both descending and ascending staircases, with reversals triggered by a change in the animal's response. The hearing threshold can be estimated by taking the average of the signal levels at a designated number of reversals or by noting the lowest level with a criterion number of "Yes" responses on ascending trials. Catch trials or silent control trials controls in which all electronics are switched on, but no test signal is projected may be used to control for response bias (see example audiometric study of a Hawaiian monk seal, Sect. 10.3.2.2). In addition, the time interval between signal presentations can be varied, so that the subject does not develop a pattern of responding based on predictable timing.

The Up/Down Staircase procedure can be difficult for an animal, because many trials are presented at near-threshold levels. This could affect an animal's motivation to respond.

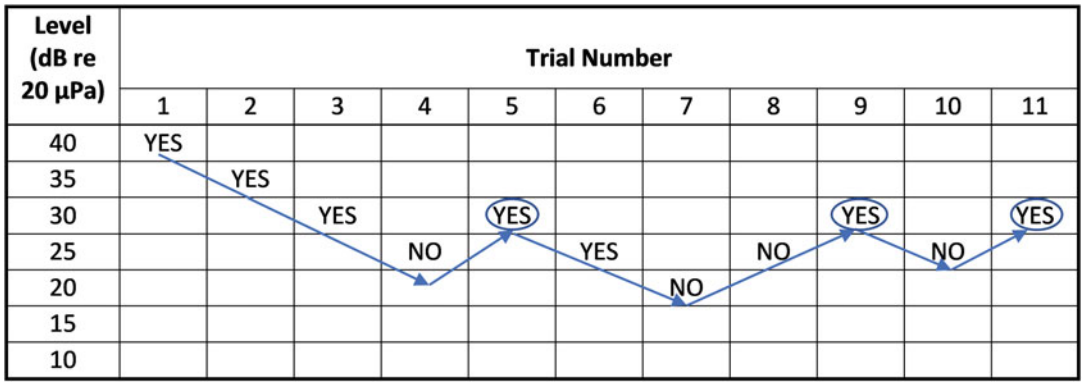



**Fig. 10.10** Example of "bracketing" a hearing threshold using the Up/Down Staircase method (Modified Method of Limits). The first signal was presented at a level that the subject easily heard ("Yes" at 40 dB re 20 µPa). Signal level was then decreased in 5-dB steps until the subject no longer signaled detection ("No" at 25 dB re 20 µPa). The change of response from "Yes" to "No" triggered the first reversal, from a descending series to an ascending one. Thereafter, each change of response triggered an immediate reversal. Signals were presented at random intervals to prevent the subject from developing a response bias based on timing. In this example, the predetermined criterion for threshold was the lowest signal level with three "Yes" responses on ascending trials (circled responses), so 30 dB re 20 µPa was the threshold for this frequency. Testing at this frequency terminated when the criterion for threshold was met

However, receiving a reward for both correct responses to signal and silent control trials helps reduce negative effects. The major advantage of the adaptive tracking method over the Method of Constant Stimuli and the Method of Limits is that fewer trials need to be conducted, resulting in a shorter test session for both the researcher and the animal subject.

### 10.3.4 Receiver Operating Characteristic (ROC) Curves

Animals, like humans, can have a bias toward a more conservative or liberal response during a hearing test (Klump et al. 1995), which could lead to underestimating or overestimating the hearing threshold, respectively. Procedures have been developed to separate response bias from actual behavioral sensitivity in psychophysical experiments. In a yes/no (audible/inaudible signal) detection task, there are four possible outcomes of each trial: (1) correct detection or hit (i.e., responding that a signal is present when it is broadcast), (2) correct rejection (i.e., responding that a signal is absent when it is not broadcast), (3) false alarm (i.e., responding that a signal is present when it is not, or indicating "yes" before the signal is broadcast), and (4) missed detection or miss (i.e., responding that a signal is absent when a signal is broadcast or failing to respond). The four response choices of an animal in a behavioral hearing test are illustrated in Fig. 10.11.

Response bias can be disentangled from sensory capabilities by constructing a Receiver Operating Characteristic (ROC) curve (Green and Swets 1966). Upon signal presentation, the

| | | Signal Presentation | |
|---|---|---|---|
| | | Signal Present | Signal Absent |
| Animal Response | Signal Present | Correct Detection (CD) | False Alarm (FA) |
| | Signal Absent | Missed Detection (MD) | Correct Rejection (CR) |

**Fig. 10.11** A two-by-two decision matrix relating the signal condition (signal presence versus signal absence) to the animal's possible responses (indicating signal presence versus signal absence) during audiometric tests

animal can respond either "yes" or "no" and so the probability of correct detection, P(CD), and the probability of missed detection, P(MD) add to 1: P(CD) + P(MD) = 1. Similarly, in the case of no signal presented, the probabilities of false alarm, P(FA), and correct rejection, P(CR), add to 1: P(FA) + P(CR) = 1. In other words, the probabilities computed from the animal responses in Fig. 10.11 are not all independent. In the ROC plot, therefore, two independent probabilities are plotted against each other: P(CD) versus P(FA). As illustrated in Fig. 10.12a, the major diagonal line marks all the points at which P(CD) = P(FA), which would be expected if the subject were making random choices or simply guessing. Below this line, the animal would perform worse than by chance; i.e., the animal would be making deliberate mistakes. The minor diagonal corresponds to P(CD) + P(FA) = 1 and so represents neutral response bias, with responses falling to the left of the line indicating a conservative response bias (i.e., low false alarm probability) and to the right a liberal response bias (i.e., high false alarm probability). The best possible performance is at the point (0|1), where the animal detects all signals and does not report any false alarms. Actual results from a beluga whale (Fig. 10.12b) detecting played-back beluga calls in icebreaker noise are shown in Fig. 10.12c. At decreasing signal-to-noise ratio (from 0 to −30 dB), the animal's hit rate decreased (i.e., decreasing P(CD)). False alarms were only made at low signal-to-noise ratio (−24 dB) indicating an overall conservative response bias. Data are based on the study by Erbe and Farmer (1998); see Fig. 10.7 for a photo of the training setup.

The bias of the animal in these hearing tests can be manipulated by changing the reinforcement regimen. If the possible responses from Fig. 10.11 are differently rewarded (e.g., positive reinforcement for the two correct responses and negative reinforcement for the two false responses), then the animal will aim to maximize the percentage of correct responses. If the four responses are all differently rewarded, then the perceived values and risks will influence the animal's response. For example, in a study with an Arctic fox (*Vulpes lagopus*; Stansbury et al. 2014), correct detections and correct rejections were rewarded with 3–4 pieces of kibble. When the animal missed a signal, it was rewarded with 1 piece of kibble. False alarms resulted in a 2–3 s time-out, after which the animal was restationed for the next trial. By rewarding misses (i.e., one of the two false responses) and with only false alarms receiving no food but instead a time-out, the animal was conditioned to avoid false alarms

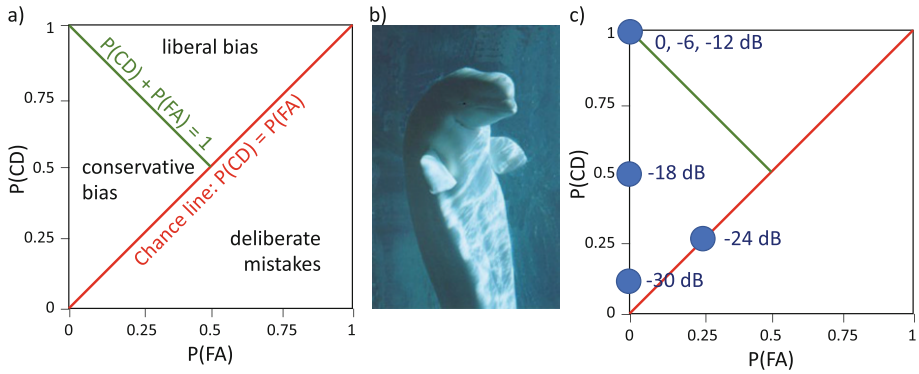



**Fig. 10.12** (**a**) Receiver Operating Characteristic (ROC) plot showing the lines and areas relating the probability of correct detection, P(CD), and the probability of false alarm, P(FA). (**b**) Photo of a beluga whale at Vancouver Aquarium. (**c**) ROC plot of this animal's performance when presented with a beluga call mixed into icebreaker noise at signal-to-noise ratios of 0, −6, −12, −18, −24, and −30 dB. The animal was trained to indicate whenever it heard the call in the noise. The animal's performance decreased with decreasing signal-to-noise ratio. The animal adopted a very conservative response bias (Erbe and Farmer 1998)

but accept misses. The reinforcement regimen directly influenced the animal's conservative bias. Similar conditioning likely happened with the beluga whale (Erbe and Farmer 1998). After the animal stationed, a sound was played randomly within a 30-s period. The animal indicated a detection (of the beluga call mixed into ice-breaker noise) by breaking from the station. If the animal did not detect a call, it held station for the full 30 s. Correct detections were rewarded with fish within 2 s. False alarms received a time-out. A "no" response received a delayed (by up to 30 s) fish reward; these would have correct rejections (i.e., signal absent trials) and missed detections (i.e., signal present trials, but under the assumption that the signal was too quiet to be detected). Effectively, the animal thus also received a reward (albeit delayed) for missed detections, even if the signal was above threshold on some of the trials. Not knowing in advance what the animal's hearing threshold is, it is impossible to tell whether the animal truly did not hear the signal when it indicated "no" to a low-level signal-present trial.

An even greater benefit of ROC analysis is realized by measuring actual ROC curves (rather than settling for scatter plots of data as in Fig. 10.12c). To do that, the animal's bias needs to be actively manipulated using reinforcement. For example, the beluga experiment could be redone with the same animal, but instead of rewarding both correct responses with one fish, the animal might be given 3 fishes for a correct detection and only 1 fish for a correct rejection. The animal might begin to favor the "yes" response, exhibiting a more liberal response bias. So, rather than having just one data point at say −12 dB signal-to-noise ratio, we would get a curve for −12 dB, with the points along the curve corresponding to the same sensitivity (hence also called isosensitivity curve) but to different biases, which were driven by the different reinforcement regimen. This is exactly what was done by Schusterman et al. (1975) with a California sea lion (*Zalophus californianus*) and a bottlenose dolphin *(Tursiops truncatus),* yielding actual ROC curves. Other ways of actively changing the bias include changing the percentage of catch trials (whereby fewer catch trials render the animal more liberal; Schusterman and Johnson 1975) or even changing the probability of handing out a reward (i.e., not all correct trials are rewarded all the time; Schusterman 1976). The resulting ROC curves then allow the separation of the animal's actual sensitivity from its bias (Green and Swets 1966; Au 1993), but much more experimental time is needed to collect all these data.

## 10.4 Physiological Methods for Audiometric Studies on Live Animals

Behavioral tests of hearing can be too time-consuming to conduct, too difficult to employ because of animals' limitations in learning or performing a behavioral task, or impractical for some other reason such as animal health, disposition, or developmental status. Physiological methods offer a practical, complementary approach because they do not require training the animal and they can be completed in a relatively shorter period of time. However, because physiological methods do not require a behavioral response from the animal that indicates the sound was perceived, they are considered to be tests of "auditory function" rather than "hearing" per se. The relationship between behavioral and physiological measures of hearing is discussed later in this chapter.

As in behavioral studies, physiological studies test responses to different kinds of acoustic stimulation and must take into account ambient noise that can affect thresholds. Other factors to consider in physiological studies are body temperature and whether or not the animal is anesthetized, because these factors can affect neural thresholds, amplitudes, and latencies. Anesthesia is commonly used in physiological studies because it is difficult to keep an unanesthetized animal in a fixed position in a sound field during testing and physical restraint can be stressful. However, anesthesia can affect brain activity and severely diminish or abolish neural responses to sound (Cui et al. 2017; Kiebel et al. 2012; McFadden and Kiebel 2013; Fig. 10.13). Anesthesia can also

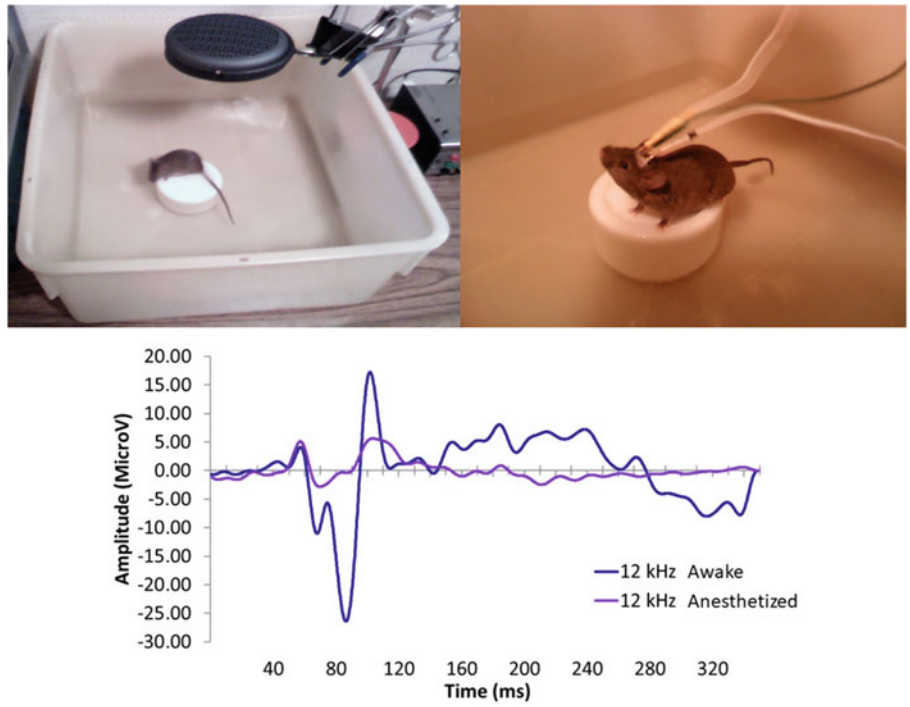



**Fig. 10.13** Top: Testing apparatus devised by Kiebel et al. (2012) for recording auditory evoked potentials from awake mice. The mice were placed on a platform (i.e., an inverted jar about 3″ in diameter) in a plastic tub containing warm water in a recording chamber. Mice were acclimated to the apparatus in daily 10-min sessions for 1–2 days prior to the first recording session. Typically, a mouse placed on the platform for the first time would enter the water and after a brief period of swimming, would climb back on the platform and remain there until removed by the researcher. In subsequent sessions, the mouse typically remained on the platform for the entire testing session (30–45 min). Stimuli were delivered from a headphone speaker placed 7″ above the animal's head. A computer-controlled camera was used to monitor the mouse, and recording was manually paused when the animal groomed or became active. Bottom: Auditory evoked responses recorded from a mouse while it was awake and then again after it had been anesthetized. The waveforms are responses to 12 kHz tones at 90 dB re 20 μPa, averaged across 100 artifact-free trials in each condition

impair thermoregulation, resulting in changes in body temperature that can be countered by placing the animal on a heating pad during testing. When brain responses must be obtained from awake animals (see Fig. 10.13), electrical artifacts created by movements during exploration or grooming can be problematic, and many trials may be required to achieve acceptable signal-to-noise ratios.

### 10.4.1 Otoacoustic Emission Methods

Otoacoustic emissions (OAEs) are sounds generated by hair cells in the inner ear, either in the absence of acoustic stimulation (spontaneous otoacoustic emissions) or in response to acoustic stimulation (transient otoacoustic emissions, TOAEs, elicited by a single tone or click; and distortion product otoacoustic emissions, DPOAEs, elicited by two primary tones, $f_1$ and $f_2$). OAEs reflect nonlinear processing in the inner ear and occur due to the action of a "cochlear amplifier," which functions to increase sensitivity to low-level sounds. Moreover, they are frequency-specific and so will emerge at those frequencies where hearing is near normal (Kemp 2002). DPOAE testing has become popular as a rapid, non-invasive way to assess the functional integrity of hair cells in a wide variety of species,

including frogs, lizards, birds, and mammals (Manley 2001). DPOAEs are abolished by loss or dysfunction of outer hair cells, and also by middle ear dysfunction that prevents retrograde transmission of acoustic energy from the cochlea to the ear canal. It is important to recognize, however, that the absence of OAEs is not necessarily evidence of outer hair cell dysfunction, because OAEs are not recordable from all normal ears. The technique is not very useful for pinnipeds because their stapedial reflex shuts down the auditory meatus as an adaptation for diving.

DPOAE tests in mammals typically use a probe assembly that is inserted into the external auditory meatus to form a closed acoustic system. For animals lacking ear canals (e.g., fishes, frogs, reptiles, and birds), the probe tip is placed inside a plastic tube that is then coupled to the animal's ear using silicone grease or Vaseline to seal any gaps (Bergevin et al. 2008). The probe tip contains a very sensitive external microphone and tubes from two external sound sources (Fig. 10.14). Two primary test tones, $f_1$ and a higher frequency tone $f_2$, are generated by separate channels of a sound-generating system and




**Fig. 10.14** A commercially available low-noise microphone with two external sound sources. The probe tip containing the microphone and sound tubes is covered with a foam or plastic ear tip and inserted into the ear canal to form a closed acoustic system. For animals without ear canals, the probe can be inserted into a plastic tube that is then sealed in place against the ear of the animal

presented through the sound tubes, and the sound in the ear canal is sampled by the microphone for a fixed period of time. The output of the microphone is filtered, digitized, averaged over a number of trials, and then analyzed using a computerized signal-analysis system. A normal inner ear will generate several nonlinear distortion products that will be propagated in a reverse direction back through the middle ear and into the ear canal (when present). When this occurs, spectrum analysis of the sound recorded by the microphone will show not only the original $f_1$ and $f_2$ tones that were delivered to the ear, but also several new tones that were generated as nonlinear distortion products. The largest distortion product is the cubic DPOAE, with a frequency equal to $2f_1 - f_2$. For example, if $f_1 = 1000$ Hz and $f_2 = 1200$ Hz, then the cochlea will generate a cubic DPOAE at 800 Hz. Because $2f_1 - f_2$ is the largest DPOAE produced (typically 30–40 dB re 20 μPa below the level of the primary tones) and is less variable than other distortion products, it is typically the only one reported in animal studies.

The frequency ratio $f_2{:}f_1$ of the primary tones, the level of the higher-frequency primary tone $L_2$, and the difference between the levels of the two primary tones $L_1 - L_2$ are selected to maximize the amplitude of the cubic DPOAE in the ear canal. These parameters are species-specific and must be determined empirically. For all combinations of stimulus parameters ($f_2{:}f_1$, $L_2$ and $L_1 - L_2$), the amplitude of the cubic DPOAE increases as the level of the primary tones increases until it saturates. DPOAEs can be difficult to measure at low frequencies due to masking by low-frequency ambient sounds in the ear canal (i.e., high noise-floor levels occur at low frequencies). But it is possible to measure low-frequency DPOAEs if great care is taken to ensure deep insertion and a good seal of the probe assembly in the ear canal.

Shaffer and Long (2004) measured low-frequency DPOAEs in two species of kangaroo rats to test the hypothesis that a large foot-drumming species (*Dipodomys spectabilis*) has better low-frequency sensitivity than a small foot-drumming species (*D. merriami*). In both species, DPOAEs were generated rated at low

frequencies between 225 and 900 Hz. DPOAE amplitudes were greater in the larger kangaroo rat species compared to the smaller species. Additionally, the authors found good correspondence between DPOAE amplitudes, behavioral hearing thresholds, and electrophysiological hearing thresholds in *D. merriami*. This suggests that DPOAE amplitudes are good estimates of hearing sensitivity.

### 10.4.2 Auditory Evoked-Potential and Auditory Brainstem Response Methods

Auditory evoked-potential (AEP) methods record stimulus-evoked electrical activity at various levels of the auditory nervous system. Hair cells and neurons in the auditory system function by generating electrical potentials in response to sounds, and measurements of these stimulus-evoked potentials can provide information about the functional state of the inner ear, auditory nerve, central auditory nuclei, and their fiber pathways (Salvi et al. 2000; McFadden 2007).

There are many ways of classifying AEPs. Common classifications are based on: (1) the region involved in the generation of the response (e.g., cochlea, brainstem, thalamus, or cortex), (2) the latency of the response (i.e., short-, middle-, and long-latency potentials reflecting generation by neural elements at progressively higher regions of the auditory system), (3) electrode placement (invasive near-field recordings made with an electrode inserted into an auditory nucleus versus noninvasive far-field recordings made from electrodes placed on the scalp), (4) the type of electrode used (high-impedance microelectrodes for recording potentials from individual cells versus low-impedance surface or needle electrodes for recording activity from large groups of neurons from the scalp), and (5) the size of the cellular population contributing to the response (e.g., local field potentials reflecting the extracellular electrical activity of a discrete group of neurons versus gross potentials generated by large populations of cells such as those recorded from scalp electrodes).

Electrical potentials generated by the cochlea and auditory nerve include the cochlear microphonic potential (CM potential) generated by outer hair cells, the summating potential (SP) generated primarily by inner hair cells, and the compound action potential (CAP) generated by the synchronous depolarization of auditory nerve fibers. AEPs generated by the auditory nerve and neurons in the auditory brainstem (i.e., cochlear nucleus, superior olive, lateral lemniscus, and inferior colliculus) contribute to the short-latency scalp-recorded auditory brainstem response (ABR). AEPs recorded from electrodes implanted into the auditory midbrain of mammals are referred to as inferior colliculus evoked potentials (IC-EVPs). AEPs generated by forebrain regions (thalamus and cortex) include long-latency potentials recorded from electrodes implanted into the brain or from surface electrodes.

AEP methods share a number of common procedures. Stimuli can be presented using the same paradigms discussed in Sect. 10.3.3 (Method of Constant Stimuli, Method of Limits, Up/Down Staircase method) with the criterion for threshold being an electrophysiological, rather than a behavioral, response. Responses are recorded and averaged over a number of trials (e.g., 50–2000 trials); the number of trials depends on the size of the response relative to background electrical noise (i.e., the signal-to-noise ratio). They are typically quantified in terms of response amplitude (e.g., peak-to-peak voltage or peak voltage relative to a baseline voltage level) and latency (i.e., the lag-time between the onset of the stimulus and a defined portion of the response). Threshold is variously defined as the lowest stimulus level that elicits a detectable physiological response, the lowest level at which a peak replicates, the midpoint between the level at which a response replicates and the next lower level at which it does not, or the sound pressure level at which the amplitude of a particular peak reaches a criterion voltage level. Other parameters that are commonly measured from AEP waveforms include peak amplitudes, peak latencies, and in the case of the ABR, interpeak intervals (i.e., time between different peaks,

reflecting neural conduction time). Results are summarized as input-output functions that show response magnitude or latency as a function of stimulus level, or as an audiogram, showing threshold as a function of stimulus frequency.

Because the ABR is an onset response that requires synchronous activity of an ensemble of neural elements, stimuli with very short rise/fall times are most effective. Clicks, which are brief (e.g., 5–100 μs) and therefore spectrally broad, often are used as stimuli, particularly for screening of auditory function. Pure tones with a rapid onset are preferred when more frequency-specific information is required, as for testing the frequency range of hearing. Sinusoidal amplitude modulated tones provide even greater frequency specificity.

At high stimulus levels that are clearly audible to an animal, several characteristic peaks are typically present in the response waveform, with latencies that correspond to their progressively higher anatomical sites of generation. ABRs from mammals typically have five prominent peaks (Fig. 10.15). The first peak of the waveform has a cochlear origin, reflecting the summed synchronous neural activity from the peripheral portion of the auditory nerve, and the second peak most likely reflects neural activity from the central portion of the auditory nerve at the level of the cochlear nucleus. Subsequent peaks are generated by brainstem regions between the cochlear nucleus and the lateral lemniscus or inferior colliculus. In all species studied, peak amplitudes of the ABR increase and latencies decrease as the stimulus level increases (Fig. 10.15). The rate of stimulus presentation can influence response amplitudes and thresholds. Data acquisition time is shortened by using a rapid signal presentation rate, but there is a cost in terms of response size, with high signal rates resulting in decreased peak amplitudes in the response waveform and increased response latencies.

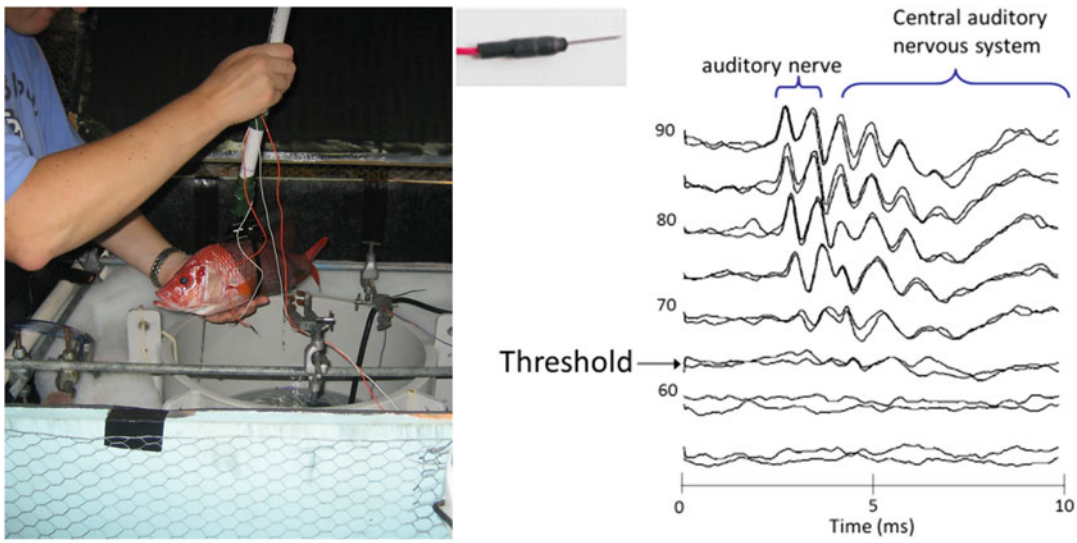



**Fig. 10.15** Left: Photo of a squirrelfish (*Sargocentron* sp.) with subcutaneous electrodes about to undergo ABR testing. Photo courtesy of Rob McCauley, Centre for Marine Science and Technology, Curtin University. Right: ABR waveforms obtained from an anesthetized C57BL/6J mouse. Needle electrodes (pictured at top left) were inserted under the skin at the top of the head (active), behind the right ear (reference), and at the base of the tail (ground). Two waveforms were collected at each stimulus level, in 5-dB steps from 90 to 55 dB re 20 μPa. Threshold, defined as the lowest level with a repeatable response, was 65 dB re 20 μPa for this frequency. The first two peaks of the ABR (short bracket) show activity from the auditory nerve, whereas the subsequent peaks (long bracket) arise from successively more rostral regions of the central auditory nervous system. Note the decrease in peak amplitude and increase in peak latency with decreasing stimulus level, typical of ABR waveforms

Preparation of animals for ABR testing is minimal. Typically, the animal is restrained or sedated or anesthetized to keep it still during the recording session. Aquatic animals under human care can be trained to remain still at a station (e.g., in a hoop) and are maintained at a good ambient water temperature in a pool. Terrestrial animals are placed on a heating pad to maintain normal body temperature. Electrodes for recording electrical activity are then applied. For most animals, the electrodes are low-impedance needle electrodes that are inserted under the skin; however, other types of electrodes, such as surface electrodes and suction-cup electrodes that attach to the surface of the head (Fig. 10.16) are suitable as well. One electrode, termed the active, non-inverting, or positive electrode, is placed at the vertex (upper surface of the head, along the midline, and between the ears) and another, termed the reference, inverting, or negative electrode, is placed behind the pinna or in another relatively neutral region of the head. A third electrode, which serves as a ground, is placed in the pool water or in a non-neural site on the animal (e.g., beneath the skin of the neck, back, or leg).

One advantage of ABRs is that it requires less time to collect a complete set of data (often 1 h or less to obtain a complete audiogram from an anesthetized animal), as compared to the weeks or months needed to train an animal for compiling behavioral audiograms. In addition, ABR testing is practical to use in studies requiring many animals and multiple measurements (e.g., before and after a treatment is applied), and for testing young animals in developmental studies. For example, McFadden et al. (1996) used ABR methods to study the ontogeny of auditory function in the Mongolian gerbil and identified three phases of development based on frequency-threshold curves. ABRs were elicited by intense stimuli in the low- and mid-frequency range as early as 10 post-natal days (pnd) in a small proportion of animals. By 16 pnd, all gerbils were responding reliably to tones between 125 Hz and 32 kHz, similar to adult animals.

ABR testing has become the AEP method of choice for audiometric testing in a wide range of species. In particular, ABRs are useful for estimating hearing capabilities of animals that are difficult to test using other methods. For example, Hu et al. (2009) used ABR recordings to determine hearing of cephalopods: the oval squid (*Sepiotheuthis lessoniana*) and the common octopus (*Octopus vulgaris*). Each cephalopod

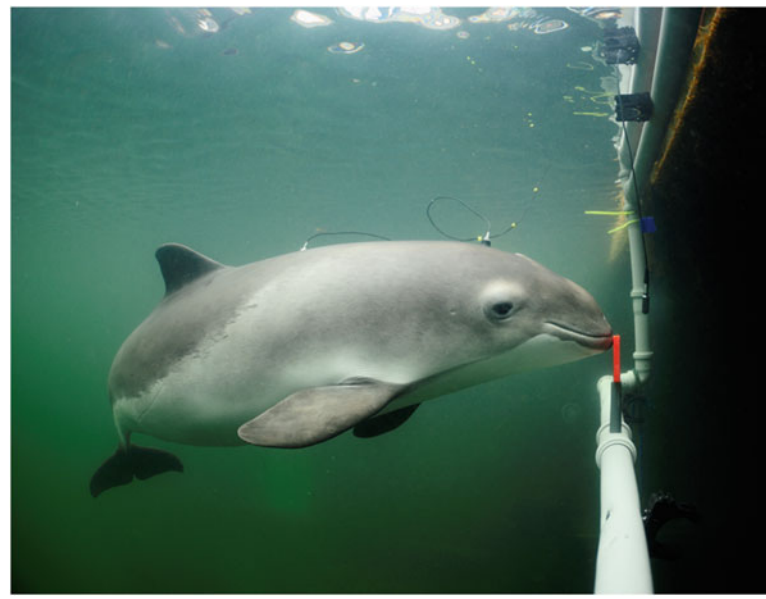


**Fig. 10.16** Photo of a harbor porpoise (*Phocoena phocoena*) stationing during an ABR test of its hearing at Fjord & Bælt Denmark. The recording electrodes, attached to the animal's head and back using suction cups, measure small electrical voltages produced by the brain in response to acoustic stimulation. Photo courtesy of Solvin Zankl, Fjord & Bælt and the Marine Biological Research Center, University of Southern Denmark, Kerteminde, Denmark

was anesthetized and then transferred to a holder inside a plastic tub filled with seawater. Teflon-coated silver needle electrodes were inserted on the head between the eyes (non-inverting) and on the mantle (inverting) and a wire was placed in the tub to serve as the ground. In both cephalopods, the ABR had only one prominent peak. The resulting ABR audiogram showed that the squid responded to a wider frequency range (400–1500 Hz vs. 400–1000 Hz) and had significantly lower thresholds at 600 Hz (its frequency of best sensitivity) compared to the octopus.

Comparisons of ABR audiograms can show the effects of factors such as age, noise exposure, drug treatment, and genetic mutations. The ABR audiograms shown in Fig. 10.17, for example, show the effects of an induced genetic mutation of the gene that codes for the copper-zinc form of superoxide dismutase (SOD1) on auditory sensitivity in mice. SOD1, an enzyme found in the cytosol of all cells, serves as a first line of defense against oxidative damage and has been implicated in numerous degenerative disorders and age-related hearing loss (McFadden et al. 2001a, b). For example, hearing thresholds of aged (13-month-old), wild type (WT) mice with normal levels of SOD1 are lower at all four tested frequencies than those of SOD1-deficient littermates. SOD1 deficiency had a greater effect on thresholds at 16 and 32 kHz than at lower frequencies (8 and 4 kHz).

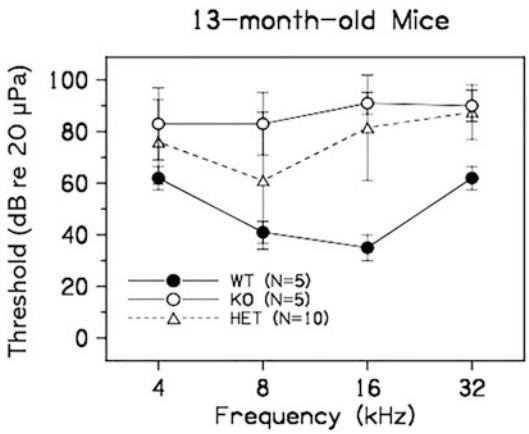



**Fig. 10.17** Average ABR thresholds (dB re 20 µPa) from aged mice with normal levels of SOD1 enzyme (WT) compared to thresholds from littermates missing 50% (HET) or 100% (KO) of SOD1 due to genetic manipulation of the copper-zinc superoxide dismutase gene. WT = wildtype mice (with two normal gene alleles and normal levels of SOD1); HET = heterozygous knockout mice (with one abnormal allele, resulting in 50% reduction of SOD1); KO = homozygous knockout mice (with two abnormal alleles, resulting in complete elimination of SOD1)

### 10.4.3 Comparison of Behavioral and Physiological Audiograms

It is important to compare data obtained from physiological and behavioral methods to determine their reliability and validity. Even in the same species, experiments might use different stimulus presentation paradigms and different threshold criteria, making direct comparisons of results difficult. Although ABR and behavioral audiograms in the same species can have the same overall shape and similar frequencies of best hearing sensitivity, actual thresholds may differ considerably (Fig. 10.18). Some authors argue that these audiograms should not be considered equivalent (Sisneros et al. 2016). Ladich and Fay (2013) compiled AEP and behavioral audiograms of goldfish collected in different studies in different laboratories. They found that, at frequencies below 1000 Hz, median ABR thresholds were about 10 dB higher than behavioral thresholds, while at higher frequencies, ABR thresholds were lower than behavioral thresholds.

Schlundt et al. (2007) quantified differences in audiograms recorded from bottlenose dolphins in a variety of underwater test conditions (in a quiet pool and in a noisy bay). AEPs were recorded using a transducer embedded in a suction cup on the jawbone. In behavioral tests, the dolphins were conditioned by the trainer's whistle to respond when the same tone was heard. Thresholds measured using the two techniques were very similar, although there was less variability in behavioral data.

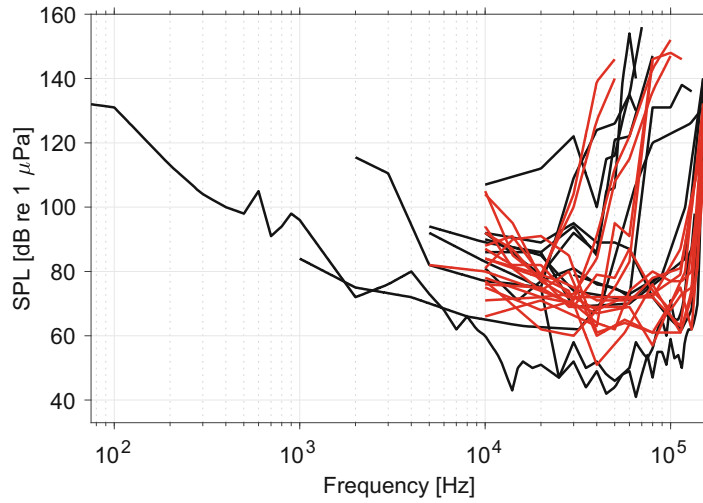



**Fig. 10.18** Comparison of underwater hearing thresholds of individual bottlenose dolphins collected by behavioral (black) versus ABR (red) methods. Data from Johnson (1966), Popov and Supin (1990), Brill et al. (2001), Houser and Finneran (2006), Finneran et al. (2008), Finneran et al. (2011)

## 10.5 Other Audiometric Measurements

Other crucial aspects of hearing can be examined using variations on the basic audiometric methods outlined above. These include frequency discrimination, intensity discrimination, equal-loudness functions, frequency selectivity (e.g., critical ratios, critical bandwidths, and psychophysical tuning curves), masking (i.e., forward, backward, and simultaneous), duration discrimination, stimulus generalization, and directional hearing (i.e., sound localization). All of these aspects of hearing have been studied in a wide range of vertebrate species. Fay (1988) compiled results of behavioral experiments from a large number of different species. Klump et al. (1995) provided complete descriptions of behavioral methods that have been developed for these kinds of experiments. Selected examples of these experiments are discussed briefly below. It is important to note that physiological techniques can also be used to obtain information on these other aspects of hearing, but that again, estimates of sensitivity may differ.

### 10.5.1 Frequency and Intensity Discrimination

Frequency and intensity discrimination experiments measure the smallest difference in frequency or intensity that an animal can detect—called the just noticeable difference (jnd) or the difference limen (DL). To measure a frequency DL using behavioral methods, the animal is trained to detect a frequency difference ($\Delta F$) between two test tones. In a typical paradigm, the animal is presented with a constant stimulus (i.e., a tone burst of one frequency) that sometimes changes in frequency, and the animal is trained to respond when it perceives a frequency change. The smallest frequency difference that the animal can perceive reliably, according to some set criterion, is the jnd or DL. Because the animal is discriminating between two frequencies, a common criterion for threshold is 75% correct, which is midway between chance and perfect performance.

Heffner and Heffner (1982) measured frequency DLs in an Indian elephant (*Elephas maximus indicus*) housed in a zoo. The elephant was trained to press one of two response buttons on a panel with its trunk upon hearing a sound. When she heard a train of tone pulses with all the same frequencies, then the correct response was to press the left button. When she heard a train of tone pulses that alternated between two different frequencies, then the correct response was to press the right button. Correct responses were rewarded with a fruit-flavored sugar solution. The DL was determined by reducing the frequency difference between the tones in the two

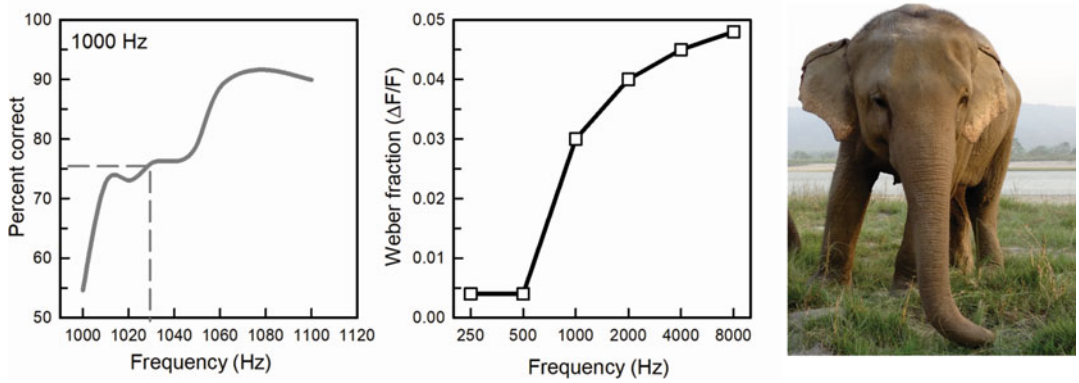



**Fig. 10.19** Psychometric function at a tone frequency of 1000 Hz (left) and a graph of the Weber fraction across frequency (middle) collected from an Indian elephant (right). Left: A psychometric function showing percent correct detection of a frequency difference between two tones. The base frequency is 1000 Hz, and frequency differences range from 20 to 100 Hz. The solid gray line shows the elephant's performance and the dashed gray line shows the 75% correct criterion for the frequency DL. At 1000 Hz, the frequency difference limen is 30 Hz. Middle: The Weber fraction (ΔF/F) increases with frequency. The Weber fraction is low at frequencies of 250 and 500 Hz, indicating good ability to discriminate frequency differences, and increases at higher frequencies, indicating poorer acuity. Data collected by Heffner and Heffner (1982). Image of the elephant from Evelyn Fuchs, University of Vienna

types of pulse trains, until the animal no longer detected the difference reliably. A psychometric function for a tone frequency of 1000 Hz, a frequency of best sensitivity for the elephant, is plotted in Fig. 10.19. The 75% correct discrimination threshold is at 1030 Hz, giving a DL or 30 Hz. The DLs calculated from psychometric functions at different tone frequencies are plotted in Fig. 10.19 as the Weber fraction (ΔF/F) the ratio of the DL to the test frequency. The Weber fraction increases with frequency, showing that the ability to discriminate differences in tone frequency becomes absolutely worse with increases in frequency. Changes in the Weber fraction with tone frequency have implications for understanding how frequency is coded in the nervous system across different species.

The psychometric function illustrated in Fig. 10.19 is based on actual data points. Some investigators use a statistical procedure called Probit Analysis to find the best-fitting regression line through the data points, and then base the estimate of the DL from that regression (Levitt 1970). The center of the best-fitting regression line can then be taken as the most probable threshold value. Probit analysis is useful because it provides a standard error for the hearing threshold values.

Intensity DLs are estimated using similar procedures as used for estimating frequency DLs, except that tone frequency is kept constant while tone intensity is varied. Difference limens are also commonly measured for noise. These measurements are useful for estimating a species' dynamic range of hearing, the intensity range over which changes in sound levels can be perceived. Determining an animal's sensitivity to the depth of amplitude modulation in a sound and the ability to detect a short, silent gap between two sounds is also a problem of intensity discrimination.

### 10.5.2 Frequency Selectivity

Frequency selectivity refers to the perceptual ability to discriminate two simultaneous signals of different frequency (e.g., a signal against noise). Behavioral measures of frequency selectivity are used to estimate the width of internal auditory filters (i.e., the physical space including number of hair cells and portion of the sensory epithelia)

devoted to a particular frequency or frequency range along the basilar membrane or sensory surface in the inner ear. Thus, behavioral measures of frequency selectivity provide an estimate of the resolving power of the ear. Physiological techniques are used to provide a more direct measurement. Auditory filters are often thought of as a series of contiguous bands of frequency in which the auditory system analyses incoming sound, and sounds of different frequencies are processed in different filters (i.e., independently of one another) without mutual interference. For ease of modeling, auditory filters often are assumed to be rectangular in shape. For very sharp frequency selectivity, hence good ability to separate signals from noise, auditory filters should be narrow. Wide auditory filters are susceptible to greater masking. Different measures of frequency selectivity exist (e.g., Fletcher critical bands, critical bandwidths, equivalent rectangular bandwidths, etc.; Fig. 10.20).

#### 10.5.2.1 Critical Ratio

The critical ratio (CR) can be thought of as the minimum signal-to-noise ratio for detecting a tone against a background of broadband masking noise. It is defined as the mean-square sound pressure of a narrowband signal (e.g., a tone) divided by the mean-square sound pressure spectral density of the masking noise at a level, where the signal is just detectable (ISO 18405:2017). 'Just detectable' again refers to a specified fraction of trials in behavioral experiments. The CR is typically expressed as a level-quantity in dB with a reference value of 1 Hz. Therefore, the CR can also be computed as the difference between the sound pressure level of the signal and the power spectral density level of the noise—at detection threshold. To measure the CR, the levels of signal (or noise) are changed. As with measuring audiograms, the CR can be measured behaviorally using the Method of Constant Stimuli, the Method of Limits, or the Up/Down Staircase

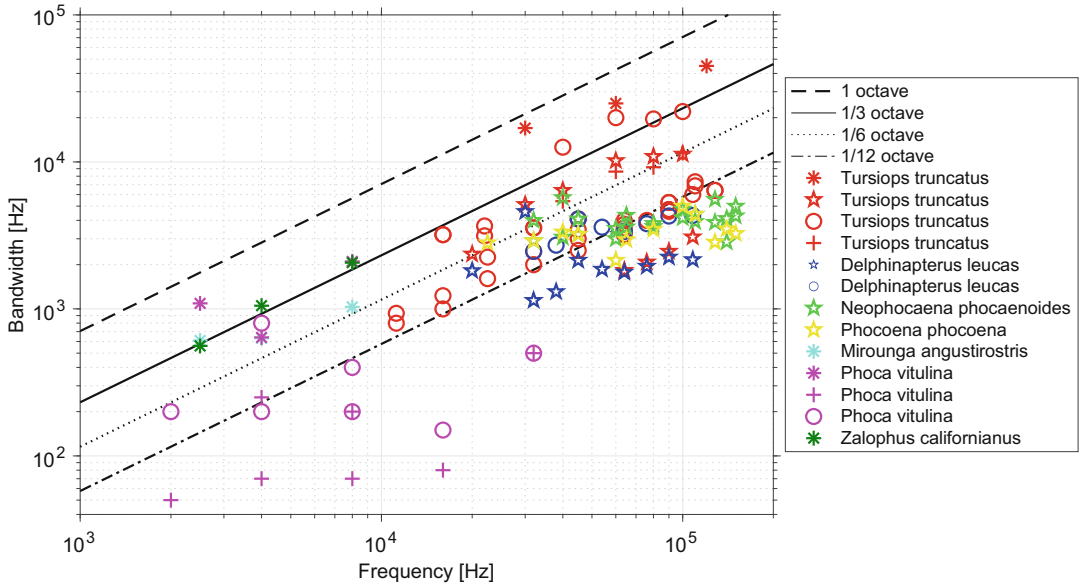



**Fig. 10.20** Graph of frequency selectivity in marine mammals. *: Critical bandwidths. ★: Equivalent rectangular bandwidths. +: 3-dB bandwidths. O: 10-dB bandwidths. Some of these data were collected behaviorally, others electrophysiologically. For pinnipeds, both in-air and underwater measurements are shown (Erbe et al. 2016). © Erbe et al. 2016; https://www.sciencedirect.com/science/article/pii/S0025326X15302125. Licensed under CC BY 4.0; https://creativecommons.org/licenses/by/4.0/

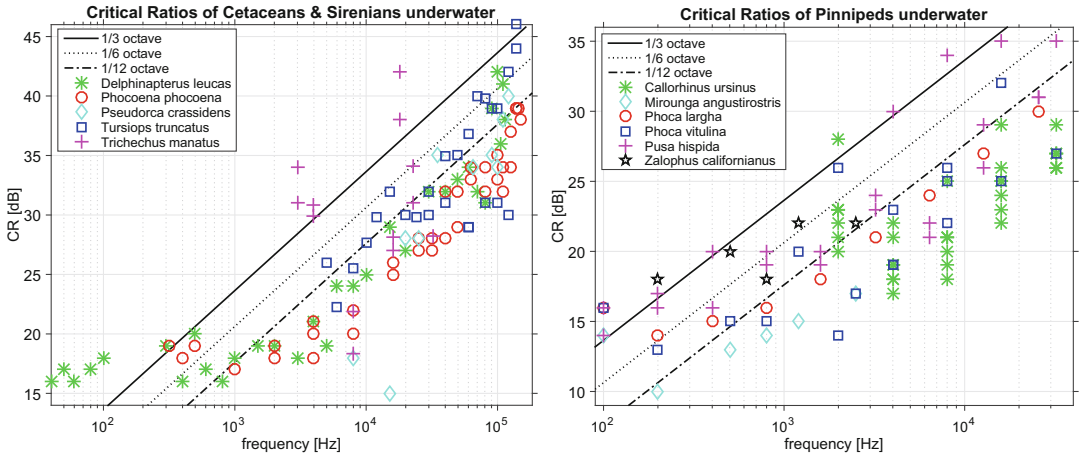



**Fig. 10.21** Graphs of critical ratios in dB re 1 Hz of marine mammals under water (Erbe et al. 2016). Fractional octave lines are shown for comparison. © Erbe et al. 2016; https://www.sciencedirect.com/science/article/pii/S0025326X15302125. Licensed under CC BY 4.0; https://creativecommons.org/licenses/by/4.0/

method. The CR can also be measured electrophysiologically.

CR measurements are relatively easy to obtain and are thus available for a number of species. In the horseshoe bat (*Rhinolophus ferrumequinum*) and in the green treefrog, for example, CRs are lowest, implying sharper filters, at the spectral peaks within this species' echolocation and advertisement calls, respectively (Long 1977; Moss and Simmons 1986). In many other species, CRs gradually increase with tone frequency (e.g., Fay 1988; Erbe et al. 2016). In the absence of CR data, 1/3 octave bands are often used (in particular in the noise impact assessment literature). While this is a good approximation in birds (e.g., Dooling and Blumenrath 2013), in several species, 1/3 octave bands overestimate CRs at some frequencies (Fig. 10.21).

The CR is often taken as an estimate of the width of the auditory filters. In this case, it should be referred to as the Fletcher critical band (ANSI/ASA S3.20-2015).[2] If CR is in dB re 1 Hz, then the Fletcher critical band is computed as $10^{CR/10}$. The Fletcher critical band is an indirect estimate of the size of the auditory filter. It is a good approximation in some bird species (Langemann et al. 1995) but in many other species differs from a more direct measure, the critical bandwidth.

### 10.5.2.2 Critical Bandwidth

The critical bandwidth (CB) refers to a band of frequencies within which sound at any frequency can interfere with sound at the center frequency (ANSI/ASA S3.20-2015; ISO 18405: 2017). The critical bandwidth is typically measured in noise-widening experiments. The listener tries to detect a tone at the center of a band of masking noise. As the noise band is widened, the level of the tone has to increase for it to remain audible. There comes a bandwidth, at which the width of the masking noise band no longer affects the level of the tone at detection threshold. This is the critical bandwidth. The difference between a CR and a CB experiment thus is that the listener has to detect a tone in broadband masking noise in the former and in noise of variable (increasing) bandwidth in the latter. CBs are time-consuming to collect, because they require determining masked thresholds at each tone frequency at many different noise bandwidths. For this reason, measurements of CB are available for fewer species than are measurements of CR.

[2] Acoustical Society of America, Standard Acoustical & Bioacoustical Terminology Database: https://asastandards.org/working-groups-portal/asa-standard-term-database/; accessed 7 January 2021.

#### 10.5.2.3 Psychophysical Tuning Curves

Psychophysical tuning curves are another measure of behavioral frequency selectivity. In these experiments, a tone is fixed in frequency and amplitude just above (typically, 10 dB) its absolute threshold. The animal is trained to detect the tone in the presence of a masker (either other tones or narrowband noise). The masker can be presented simultaneously with the tone (simultaneous masking), or prior to the tone (forward masking). Psychophysical tuning curves are typically V-shaped, so that as the frequency separation between the tone and the masker increases, the level of the masker required to mask the tone increases (Fig. 10.22). They are similar in shape to tuning curves of auditory nerve fibers, and so can provide non-invasive estimates of neural frequency selectivity (Serafin et al. 1982). The drawback of this technique is that it is time-consuming to conduct, so that data are available for only a few animal species.

## 10.6 Summary

Describing and quantifying the hearing capabilities of different animals is essential in bioacoustical studies. Basic features of hearing, such as the range of audibility, thresholds of hearing as a function of frequency, and the frequency range of best hearing, are easily shown on an audiogram. Hearing sensitivities are best in young, healthy animals and may decline in some animals as they age or if they are exposed to ototoxic antibiotics. Acute exposure to high-amplitude noise or long-term exposure to lower levels of noise also can temporarily or permanently reduce hearing sensitivity.

A variety of behavioral and physiological methods can be used to test hearing in live animals. The aims of a study and the characteristics of the animals should be considered carefully when selecting the appropriate audiometric methods to use. This chapter

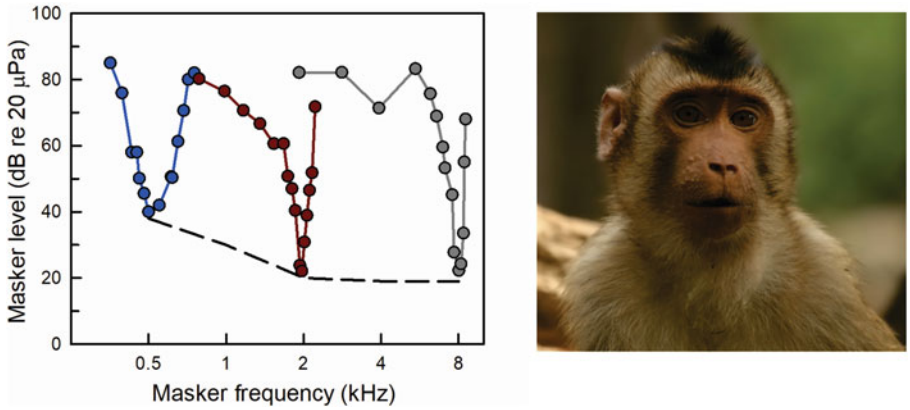



**Fig. 10.22** Psychophysical tuning curves (left) for the Pig-tailed macaque monkey (*Macaca nemestrina;* right), measured in a forward masking paradigm. Animals were trained to detect tones using positive reinforcement. Tones were presented via earphones, and the animals were seated inside a sound-attenuating chamber. Masked thresholds to probe tones (0.5, 2, and 8 kHz; blue, dark red, dark gray, respectively; x-axis) were determined using an adaptive tracking procedure and defined as the mean of eight reversal points at each frequency. Probe tones (25-ms duration) were presented at a level of 10 dB above absolute threshold. Masker tones (130-ms duration, with frequencies varying around that of the probe tone) were presented 2 ms before the onset of the probe tone. The blue, dark red, and dark gray curves show the psychophysical tuning curves plotting the level of the masker (y-axis) needed to just mask the probe tone at each masker frequency. The black dashed line shows the animals' absolute thresholds (audiogram). Data collected by Serafin et al. (1982). © Stauss, 2006; https://commons.wikimedia.org/w/index.php?curid=1733069. Licensed under CC BY-SA 3.0; https://creativecommons.org/licenses/by-sa/3.0/

described common behavioral and physiological methods, along with some of their strengths and weaknesses. Testing hearing abilities in animals is not as easy as in humans because animal subjects cannot verbally report to the researcher when a test signal is heard. Instead, animals indicate that they heard a sound by making unlearned or learned responses in behavioral studies. Thresholds based on conditioned responses are the most accurate and reliable, but conditioning procedures are not suitable for all animals or research questions. Some animals are not trainable or are unable to participate in a behavioral study due to age, health, or some other factor. Physiological methods, especially auditory brainstem response testing, can be particularly helpful in these situations. While ABR and other physiological methods provide useful information about auditory function, it is important to recognize that the results they provide are not equivalent to those from behavioral studies that assess hearing directly; thresholds obtained using physiological methods may under- or overestimate behavioral thresholds in an unpredictable manner.

Research on hearing abilities in animals has advanced beyond documenting the basic audiogram of a species. Data on frequency and intensity discrimination, sound localization, and the effects of noise on hearing in animals are current topics of study for many animal species. Information on hearing and an animal's abilities to adapt to noise can have important applications for the conservation of species in areas of high anthropogenic noise.

## References


Aihara I, Bishop PJ, Ohmer MEB, Awano H, Mizumoto T, Okuno HG, Narins PM, Hero JM (2017) Visualizing phonotactic behavior of female frogs in darkness. Sci Rep 7:10539. https://doi.org/10.1038/s41598-017-11150

American National Standards Institute (2004) Methods for manual pure-tone threshold audiometry (ANSI S3.21-2004). Acoustical Society of America, New York

American National Standards Institute (2015) Bioacoustical terminology (ANSI S3.20-2015, R 2020). Acoustical Society of America, New York

Au WWL (1993) The sonar of dolphins. Springer, New York

Awbrey FT, Thomas JA, Kastelein RA (1988) Low-frequency underwater hearing sensitivity in belugas, *Delphinapterus leucas*. J Acoust Soc Am 84(6):2273–2275

Bain DE, Kriete B, Dahlheim M (1993) Hearing abilities of killer whales (*Orcinus orca*). J Acoust Soc Am 94(3):1829–1829

Berger JI, Coomber B, Shackleton TM, Palmer AR, Wallace MN (2013) A novel behavioural approach to detecting tinnitus in the Guinea pig. J Neurosci Methods 213(2):188–195

Bergevin C, Freeman DM, Saunders JC, Shera CA (2008) Otoacoustic emissions in humans, birds, lizards, and frogs: evidence for multiple generation mechanisms. J Comp Physiol A 194(7):665–683

Bhandiwad AA, Sisneros JA (2016) Revisiting psychoacoustic methods for the assessment of fish hearing. In: Sisneros JA (ed) Fish hearing and bioacoustics: an anthology in honor of Arthur N. Popper and Richard R. Fay, vol 877. Springer, New York, pp 157–184

Bowles AE, Francine JK (1993) Effects of simulated aircraft noise on hearing, food detection, and predator avoidance behavior of the kit fox, *Vulpes macrotis*. J Acoust Soc Am 93:2378–2378

Braff DL, Geyer MA, Swerdlow NR (2001) Human studies of prepulse inhibition of startle: normal subjects, patient groups, and pharmacological studies. Psychopharmacology 156(2–3):234–258

Branstetter BK, Leger JS, Acton D, Stewart J, Houser D, Finneran JJ, Jenkins K (2017) Killer whale (*Orcinus orca*) behavioral audiograms. J Acoust Soc Am 141(4):2387–2398. https://doi.org/10.1121/1.4979116

Brill RL, Moore PWB, Dankiewicz LA (2001) Assessment of dolphin (*Tursiops truncatus*) auditory sensitivity and hearing loss using jawphones. J Acoust Soc Am 109(4):1717–1722

Cui J, Zhu B, Fang G, Smith E, Brauth SE, Tang Y (2017) Effect of the level of anesthesia on the auditory brainstem response in the Emei Music Frog (*Babina daunchina*). PLoS One 12(1):e0169449. https://doi.org/10.1371/journal.pone.0169449

Dent ML, Dooling RJ, Pierce AS (2000) Frequency discrimination in budgerigars (*Melopsittacus undulatus*): effects of tone duration and tonal context. J Acoust Soc Am 107(5):2657–2664. https://doi.org/10.1121/1.428651

Dooling RJ, Blumenrath SH (2013) Avian sound perception in noise. In: Brumm H (ed) Animal communication in noise. Springer, Heidelberg, pp 229–250. https://doi.org/10.1007/978-3-642-41494-7_8

Ehret G, Romand R (1981) Postnatal development of absolute auditory thresholds in kittens. J Comp Physiol Psychol 95(2):304–311

Erbe C, Farmer DM (1998) Masked hearing thresholds of a beluga whale (*Delphinapterus leucas*) in icebreaker noise. Deep Sea Res II Top Stud Oceanogr 45(7):1373–1388. https://doi.org/10.1016/S0967-0645(98)00027-7

Erbe C, Reichmuth C, Cunningham KC, Lucke K, Dooling RJ (2016) Communication masking in marine mammals: a review and research strategy. Mar Pollut Bull 103:15–38. https://doi.org/10.1016/j.marpolbul.2015.12.007

Fay RR (1988) Hearing in vertebrates: a psychophysics databook. Hill-Fay Associates, Winnetka, IL

Fay RR (1995) Psychoacoustical studies of the sense of hearing in goldfish using conditioned respiratory suppression. In: Klump GM, Dooling RJ, Fay RR, Stebbins WC (eds) Methods in comparative psychoacoustics. Birkhauser, Basel, pp 249–261

Fay RR, Simmons AM (1999) The sense of hearing in fishes and amphibians. In: Popper AN, Fay RR (eds) Comparative hearing: fish and amphibians. Springer, New York, pp 269–318

Finneran JJ, Houser DS, Blasko D, Hicks C, Hudson J, Osborn M (2008) Estimating bottlenose dolphin (*Tursiops truncatus*) hearing thresholds from single and multiple simultaneous auditory evoked potentials. J Acoust Soc Am 123(1):542–551

Finneran JJ, Mulsow J, Schlundt CE, Houser DS (2011) Dolphin and sea lion auditory evoked potentials in response to single and multiple swept amplitude tones. J Acoust Soc Am 130(2):1038–1048. https://doi.org/10.1121/1.3608117

Gerhardt HC (1995) Phonotaxis in female frogs and toads: execution and design of experiments. In: Klump GM, Dooling RJ, Fay RR, Stebbins WC (eds) Methods in comparative psychoacoustics. Birkhauser, Basel, pp 209–220

Green D, Swets J (1966) Signal detection theory and psychophysics. Wiley, New York. Reprinted 1974 by Krieger, Huntington, New York

Hall JD, Johnson CS (1972) Auditory thresholds of a killer whale, *Orcinus orca* Linnaeus. J Acoust Soc Am 51: 515–517

Heffner RS, Heffner HE (1982) Hearing in the elephant (*Elephas maximus*): absolute sensitivity, frequency discrimination, and sound localization. J Comp Physiol Psychol 96:926–944

Heffner RS, Heffner HE (1991) Behavioral hearing range of the chinchilla. Hear Res 52:13–16

Heffner HE, Heffner RS (2001) Behavioral assessment of hearing in mice. In: Willott JF (ed) Handbook of mouse auditory research: from behavior to molecular biology. CRC Press, Boca Raton, FL, pp 19–29

Heffner RS, Heffner HE, Masterton B (1971) Behavioral measurements of absolute and frequency-difference thresholds in Guinea pig. J Acoust Soc Am 49(6B):1888–1895

Heffner HE, Heffner RS, Contos C, Ott T (1994) Audiogram of the hooded Norway rat. Hear Res 73:244–248

Heffner RS, Koay G, Heffner HE (2014) Hearing in alpacas (*Vicugna pacos*): audiogram, localization acuity, and use of binaural locus cues. J Acoust Soc Am 135(2):778–788

Hellström P-A (1995) The relationship between sound transfer functions and hearing levels. Hear Res 88(1–2):54–60

Henderson D, Salvi RJ, Quaranta A, McFadden SL, Burkard RF (eds) (1999) Ototoxicity: basic science and clinical applications, Annals of the New York Academy of Sciences, vol 884. The New York Academy of Sciences, New York

Hoffman H, Ison JR (1980) Reflex modification in the domain of startle: I. Some empirical findings and their implications for how the nervous system processes sensory input. Psychol Rev 87:175–189

Houser DS, Finneran JJ (2006) A comparison of underwater hearing sensitivity in bottlenose dolphins (*Tursiops truncatus*) determined by electrophysiological and behavioral methods. J Acoust Soc Am 120:1713–1722

Hu MY, Yan HY, Chung W-S, Shiao J-C, Hwant PP (2009) Acoustically evoked-potentials in two cephalopods inferred using the auditory brainstem response (ABR) approach. Comp Biochem Physiol A 153(3):278–283

International Organization for Standardization (2017) Underwater acoustics—terminology (ISO 18405). Switzerland, Geneva

Jero J, Coling DE, Lalwani AK (2001) The use of Preyer's reflex in evaluation of hearing in mice. Acta Otolaryngol 121(5):585–589

Johnson CS (1966) Auditory thresholds of the bottlenose porpoise (*Tursiops truncatus*, Montagu). U.S. Naval Ordnance Test Station. Tech Publ 4178:1–28

Kastelein RA, Heu S, van der Verboom W, Jennings N, Veen J, Vander J, de Haan D (2008) Startle response of captive North Sea fish species to underwater tones between 0.1 and 64 kHz. Mar Environ Res 65(5): 369–377

Kemp DT (2002) Otoacoustic emissions, their origin in cochlear function, and use. Br Med Bull 63(1): 223–241

Kiebel EM, Sunderman MG, Leonhard JR, McFadden SL (2012) Measurement of cortical auditory event-related potentials in conscious CBA/CaJ mice. Association for Psychological Science conference, Chicago

Klump GM, Dooling RJ, Fay RR, Stebbins WC (eds) (1995) Methods in comparative psychoacoustics. Birkhauser, Basel

Koay G, Heffner RS, Heffner HE (2002) Behavioral audiograms of homozygous medJ mutant mice with sodium channel deficiency and unaffected controls. Hear Res 171:111–118

Ladich F, Fay RR (2013) Auditory evoked-potential audiometry in fish. Rev Fish Biol Fish 23:317–364

Langemann U, Klump GM, Dooling RJ (1995) Critical bands and critical-ratio bandwidth in the European starling. Hear Res 84(1–2):167–176

Levitt H (1970) Transformed up-down methods in psychoacoustics. J Acoust Soc Am 49:467–477

Long GR (1977) Masked auditory thresholds from the bat, *Rhinolophus ferrumequinum*. J Comp Physiol 116: 247–255

Manley GA (2001) Evidence for an active process and a cochlear amplifier in nonmammals. J Neurophysiol 86(2):541–549

McFadden SL (2007) Biochemical bases of hearing. In: Campbell K (ed) Pharmacology and ototoxicity for

audiologists. Thomson Delmar Learning, New York, pp 86–123
McFadden SL, Kiebel EM (2013) A parametric study of auditory event-related potentials recorded from cortex of CBA/CaJ mice. Association for Psychological Science Conference, Washington, DC
McFadden SL, Walsh EJ, McGee J (1996) Onset and development of auditory brainstem response thresholds in the Mongolian gerbil (*Meriones unguiculatus*). Hear Res 100:68–79
McFadden SL, Campo P, Quaranta N, Henderson D (1997) Age-related decline of auditory function in the chinchilla (*Chinchilla laniger*). Hear Res 111:114–126
McFadden SL, Ohlemiller KK, Ding DL, Salvi RJ (2001a) The role of superoxide dismutase in age-related and noise-induced hearing loss: clues from *Sod1* knockout mice. In: Willott JF (ed) Handbook of mouse auditory research: from behavior to molecular biology. CRC Press, Boca Raton, FL, pp 489–504
McFadden SL, Ohlemiller KK, Ding DL, Salvi RJ (2001b) The influence of superoxide dismutase and glutathione peroxidase deficiencies on noise-induced hearing loss in mice. In: Henderson D, Prasher D, Kopke R, Salvi R, Hamernik R (eds) Noise-induced hearing loss: basic mechanisms, prevention and control. NRN Publications, London, pp 3–18
McFadden SL, Zulas AL, Morgan RE (2010) Age-dependent effects of modafinil on acoustic startle and prepulse inhibition in rats. Behav Brain Res 208(1):118–123
Moss CF, Simmons AM (1986) Frequency selectivity of hearing in the green treefrog, *Hyla cinerea*. J Comp Physiol A 159:257–266
Popov VV, Supin AY (1990) Electrophysiological studies on hearing in some cetaceans and a manatee. In: Thomas JA, Kastelein RA (eds) Sensory abilities of cetaceans: laboratory and field evidence. Plenum Press, New York, pp 405–415
Salvi RJ, McFadden SL, Wang J (2000) Anatomy and physiology of the peripheral auditory system. In: Roeser RJ, Valente M, Hosford-Dunn H (eds) Audiology diagnosis. Thieme, New York, pp 19–43
Schlundt CE, Dear RL, Green L, Houser DS, Finneran JJ (2007) Simultaneously measured behavioral and electrophysiological hearing thresholds in a bottlenose dolphin (*Tursiops truncatus*). J Acoust Soc Am 122:615–622
Schusterman RJ (1976) California Sea lion underwater auditory detection and variation of reinforcement schedules. J Acoust Soc Am 59(4):997–1000. https://doi.org/10.1121/1.380928
Schusterman RJ, Johnson BW (1975) Signal probability and response bias in California Sea lions. Psych Rec 25(1):39–45. https://doi.org/10.1007/BF03394287
Schusterman RJ, Barrett B, Moore P (1975) Detection of underwater signals by a California Sea lion and a bottlenose porpoise: variation in the payoff matrix. J Acoust Soc Am 57(6):1526–1532. https://doi.org/10.1121/1.380595
Screven LA, Dent ML (2019) Perception of ultrasonic vocalizations by socially housed and isolated mice. eNeuro 6(5). https://doi.org/10.1523/ENEURO.0049-19.2019
Serafin JV, Moody DB, Stebbins WC (1982) Frequency-selectivity of the monkey's auditory system: psychophysical tuning-curves. J Acoust Soc Am 71(6):1513–1518
Shaffer LA, Long GR (2004) Low-frequency distortion product otoacoustic emissions in two species of kangaroo rats: implications for auditory sensitivity. J Comp Physiol A 190(1):55–60
Simmons AM, Moss CF (1995) Reflex modification: a tool for assessing basic auditory function in anuran amphibians. In: Dooling R, Fay R, Klump G, Stebbins W (eds) Methods in comparative psychoacoustics. Birkhauser, Basel, pp 197–208
Sisneros JA, Popper AN, Hawkins AD, Fay RR (2016) Auditory evoked-potential audiograms compared to behavioral audiograms in aquatic animals. In: Popper AN, Hawkins AD (eds) Effects of noise on aquatic life II, vol 875. Springer, New York, pp 1049–1056
Stansbury AL, Thomas JA, Stalf CE, Murphy LD, Lombardi D, Carpenter J, Mueller T (2014) Behavioral audiogram of two Arctic fox (*Alopex lagopus*). Polar Biol 37:417–422. https://doi.org/10.1007/s00300-014-1446-5
Thomas JA, Moore PWB, Withrow R, Stoermer M (1990) Underwater audiogram of a Hawaiian monk seal (*Monachus schauinslandii*). J Acoust Soc Am 87(1):417–420
Tonndorf J (1976) Relationship between the transmission characteristics of conductive system and noise-induced hearing-loss. In: Henderson D, Hamernik RP, Dosanjh DS, Mills JH (eds) Effects of noise on hearing. Raven Press, New York, pp 159–178
von Békésy G (1960) Experiments in hearing. McGraw Hill, New York
Walter M, Tziridis K, Ahlf S, Schulze H (2012) Context dependent auditory thresholds determined by brainstem audiometry and prepulse inhibition in Mongolian gerbils. Open J Acoust 2:34–49. https://doi.org/10.4236/oja.2012.21004
Welch TE, Dent ML (2011) Lateralization of acoustic signals by dichotically listening budgerigars (*Melopsittacus undulatus*). J Acoust Soc Am 130(4):2293–2301
White JR, Norris MJH, Ljungblad DK, Barton K, di Sciarra GN (1978) Auditory thresholds of two beluga whales, *Delphinapterus leucas*. HSWRI Tech Rep No 78-109 Sea World Research Institute, San Diego, CA
Willott JF (1991) Aging and the auditory system: anatomy, physiology, and psychophysics. Singular Publication Group, San Diego, CA

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



# Vibrational and Acoustic Communication in Animals

11

Rebecca Dunlop, William L. Gannon, Marthe Kiley-Worthington, Peggy S. M. Hill, Andreas Wessel, and Jeanette A. Thomas

## 11.1 Introduction

The study of animal communication, which is sometimes called zoosemiotics (as opposed to anthroposemiotics, the study of human communication), is fundamental to the areas of ethology, evolutionary biology, and animal cognition. Here, we are not so emboldened as to claim that humans are separate from other "animals." In fact, we are ordinary mammals. Therefore, other than a brief discussion of human language at the end of the chapter, we will not discuss anthroposemiotics. Instead, we highlight and discuss what much of the rest of the Kingdom does.

In *Acoustic behavior of animals* (edited by Busnel 1963, p. 751), Tembrock stated that, "the production of sounds is not a fancy of Nature, but an expression of biological needs." Moles (also in Busnel 1963), in what are believed to be the main lines of acoustic communication in animals, included a code that is received and acted upon (p. 112). Groundbreaking as this volume was, knowledge of acoustic communication in animals has come a long way since. Just 20 years later, Kroodsma (1982) published *Acoustic communication in birds*. The first volume of this multivolume publication discussed the significant advances made in recording animal signals, as well as the advancement in knowledge of the anatomy of neural and auditory structures, the physical characters of signal transmission, signaler motivation and coding, species-specific signaling, and the use of signals in behaviors such as spacing and mating (Morton 1982). The second volume (Kroodsma and Miller 1982) discussed issues of signal ontogeny, mimicry, vocal learning, and the ecological, behavioral, and genetic implications of variations within vocalizations. Other early compendiums, such as Sebeok (1977), provided an extensive summary of high-quality research studies from an expanding discipline of behavior and animal communication.

Jeanette A. Thomas (deceased) contributed to this chapter while at the Department of Biological Sciences, Western Illinois University-Quad Cities, Moline, IL, USA

R. Dunlop (✉)
School of Biological Sciences, University of Queensland, Brisbane, QLD, Australia
e-mail: r.dunlop@uq.edu.au

W. L. Gannon
Department of Biology, Museum of Southwestern Biology, and Graduate Studies, University of New Mexico, Albuquerque, NM, USA
e-mail: wgannon@unm.edu

M. Kiley-Worthington
Centre of Eco-Etho Research & Education, Cranscombe Cleave, Brendon, UK

P. S. M. Hill
Department of Biological Science, The University of Tulsa, Tulsa, OK, USA
e-mail: peggy-hill@utulsa.edu

A. Wessel
Center for Integrative Biodiversity Discovery, Museum für Naturkunde Berlin, Berlin, Germany
e-mail: andreas.wessel@mfn.berlin

© The Author(s) 2022
C. Erbe, J. A. Thomas (eds.), *Exploring Animal Behavior Through Sound: Volume 1*,
https://doi.org/10.1007/978-3-030-97540-1_11



**Fig. 11.1** Biotremology examines mechanical communication such as that produced by many insects, including planthoppers (*Apache degeeri*; common in places such as North Carolina, USA). Photo "9 *Apache degeeri* (planthopper)" by Wildreturn; https://wordpress.org/openverse/image/4323324f-25c8-408f-9b88-8c5b3ae93655/. Licensed under CC BY 2.0; https://creativecommons.org/licenses/by/2.0/

Bioacoustics is defined as the study of mechanical communication by acoustic (sound) waves. It is a widely used term when referring to animal communication. Biotremology is a relatively recent term. It was conceived to refer to communication signals that comprise substrate-borne vibrations, and which are detected as surface vibrations by specialized perception organs such as slit-sense organs in spiders, subgenual organs in insects, hair receptors, or Pacinian and Herbst corpuscles in vertebrates (Hill and Wessel 2016). Substrate-borne vibrations are sensed via, "…pressure waves traveling through … solid matter … detected via the surface vibrations they elicit or the airborne waves (sound) they induce" (Hill and Wessel 2016). Bioacoustical (sound) communication, refers to signals that are encoded in acoustic waves, and are detected using the ear. Vibrational communication has been recognized as evolutionarily older than bioacoustic communication and is much more prevalent among some animal groups (e.g., arthropods; Fig. 11.1). Therefore, researchers are also interested in how these mechanical vibrations affect behavior.

Both areas of study use similar equipment to record and analyze communication signals. However, scientists in the field of biotremology also use devices such as laser Doppler vibrometers and wavelet analysis. These function to detect faint vibrational emissions made by animals. In addition, electromagnetic transducers produce signals, and when in contact with the substrate, serve as vibration generators for artificial playback experiments.

Now, nearly 60 years later beyond Busnel's (1963) paradigm of bioacoustics, tremendous changes in recording technology and analysis have occurred. Acoustic identification of anything from birds to bats can be carried out using an iPhone, an acoustic detection application, and a bluetooth speaker or microphone!

## 11.2 The Origins of Substrate-Borne Vibrational and Acoustic Communication

Communication is the transfer of information from one animal (sender) to another animal (receiver) that can affect the current or future behavior of the receiver. In other words, communication conveys information. It is adaptive, in that a successful communication exchange enhances the survival of one or both participants. Vibrational communication has been suggested to have evolved, along with chemical

communication, concurrently with evolution of the Metazoa (all animals; Endler 2014). We know that any movement of an animal, whether in water or at the boundary between air and any type of substrate, creates vibrations that can be detected by any other organism with receptors capable of receiving and translating them. Increasing evidence also suggests that invertebrate hearing organs evolved from vibrational precursors millions of years ago (Stumpner and von Helversen 2001; Lakes-Harlan and Strauss 2014). Therefore, the discussion of origins of communication in this section is restricted to the more recently evolved acoustic communication.

The origins of acoustic communication are likely to be in nonverbal sounds made by chance as the animal moves through the environment. These sounds could be scraping, a stick breaking, footfalls, opening or flapping of wings, or scratching. They are the result of environmental disturbance, which in turn makes a sound through the air, earth, or water. By just being made, these sounds convey to others the presence of the animal, and something about what it might be doing. It is then a simple developmental step for a particular sound to become associated with a particular situation and thus carry a particular message to the recipient. Examples of nonverbal sounds are sounds from an elephant breaking sticks as it moves through the environment, a sigh, a cough, or a sneeze. Originally, these sounds may not have been made to communicate. However, sounds that provide an advantage for an individual, or a population, will be perpetuated if they enhance the fitness of the species. This, ultimately, gives them an evolutionary advantage that would reinforce further refinement of this new sensory mode.

This origin likely gave the evolutionary opening to develop specialized body parts that could produce auditory signals, in tandem with sophisticated sensory capabilities to receive them (Narins et al. 2009). One such specialized body part is the respiratory tract. Once a respiratory tract had developed in vertebrates, sounds associated with breathing could convey information to others, and so the necessary adaptations for sound generation began to develop. For example, holding the breath and then letting it out as a sigh or a cough produces various sounds. These sounds are then associated with situations being experienced by the sender, meaning this information is available to all who hear it. Presumably, it was this evolutionary process that gave rise to sound-making organs in the respiratory tract to the point where vocal communication now involves a larynx.

Ritualization is the evolutionary process by which a pattern of behavior changes to become more effective as a signal (Huxley 1966; Morris 1957). The behavior is performed in a consistent way and is either stereotyped or incomplete. Incomplete behaviors may be used for activities such as courtship. For example, a drake mallard (*Anas platyrhynchos*), when preening and displaying to a female, acts as if he is addressing a skin irritation (Morris 1956), but he may not even touch his feathers during the display. In other words, the behavior seems to be a preening behavior, but is in fact a courtship behavior. To increase the effectiveness of the ritualized signal, anatomical modifications may also have evolved. A classic example of this is the elaborate colors of the Mandarin drake (*Aix galericulata*). During the courtship of a female, the male will highlight these colors by pointing to them during incomplete, exaggerated, and stereotypical preening.

Exaggerated signal ritualization is characterized by a clear signaling behavior, such as the ears of a horse (*Equus caballus*) flattening back as a precursor signal to biting. This exaggerated ear movement has a clearer meaning than just putting the ears back. Ritualistic behavior is usually no longer tied to its original role because it has become more important for the signaler's fitness to communicate, rather than being used for its original purpose. Therefore, the signal has evolved to produce a clear message.

Signals can also evolve to become more effective by redundancy, or by emulation of another's acoustic or vibrational expression. Redundancy in animal acoustic communication is the repeated use of a signal. Vocal signals, for example, can be repeated for long periods of time, such as the continuous chorusing of frogs advertising during mating sessions. Redundancy reduces the risk



**Fig. 11.2** Emulative acoustic behavior is seen when a domestic dog (*Canis lupus familiaris*) hears a siren or other high-pitched signal. Photo "Howling white husky" by Tambako the Jaguar; https://wordpress.org/openverse/image/7d77b8d9-3dc4-4f3d-9c04-318833d1759e/. Licensed under CC BY-ND 2.0; https://creativecommons.org/licenses/by-nd/2.0/

that a signal will be missed or misinterpreted and assures that the signal is heard even when environmental conditions are poor (e.g., when there are masking sounds from the environment and/or human sources). This continual production of sounds in chorus can also sustain the state of arousal or excitement, which may be necessary for completion of the behavior.

Signal emulation is when other members of a group join in when a signal is given. An example of this is when a group of domestic dogs (*Canis lupus familiaris*) hear the high-pitched siren of an emergency vehicle. One may start to howl, and others soon join in (Fig. 11.2). When one individual calls, this often stimulates others to make the same call. Other examples include the greeting calls (trumpeting) between mother and offspring elephant (*Elephas maximus*), or the "see-saw" vocalized inspiration and expiration call and reply signals of bull cattle (*Bos taurus*). Sound emulation is also common in humans. The vocalization is copied and repeated by a recipient and can cause increased arousal in both the sender and the recipient (Kiley 1972). Animals copying new sounds, which often happens by emulation, requires vocal learning (Janik and Slater 2000).

A more complex version of this is antiphonal singing, which is an acoustic exchange between animals where they call at the same time to produce a chorus. There are benefits to this emulative calling behavior. Males that chorus, such as frogs and toads (Anura), cicadas (Cicadoidea), and humpback whales (*Megaptera novaeangliae*), may attract more females to a localized area. For example, millions of cicadas gather to mate in a forest in the eastern US, where the singing males produce loud, pure-tone sounds above 90 dB SPL (Fig. 11.3; Bennet-Clark 1998, 2000). Prairie mole cricket (*Gryllotalpa major*) males in the south-central US sing in choruses from burrows in the soil that individuals construct in aggregations. At 20-cm from the burrow entrance, the males' loud harmonic songs average 96 dB SPL (Hill 1998).

The larynx and various resonating cavities in the respiratory tract (throat, mouth, and nasal cavities that can be specialized into trunks or elongated noses) are collectively responsible for an enormous range of vocal sounds made by different species. Vocal signals have evolved to convey a great variety of messages, encompassing many meanings that can be interpreted by the recipients. The development of this messaging system becomes intricate with human language. Whether the degree of development of the young at birth (which could relate to cognitive development; Scheiber et al. 2017; Wilson-Henjum et al. 2019) influences the complexity of vocalizations and other displays are yet to be determined.

## 11.3 A Summary of Communication

Communication occurs when a signaler encodes a message in a signal, which passes through some medium (air, water, soil, plant organs, etc.), and is received, decoded, and acted upon by the receiver. The receiver's response benefits the



**Fig. 11.3** 17-year Cicada (*Magicicada* sp). Photo by the U.S. Department of Agriculture; https://www.flickr.com/photos/usdagov/8672057401/in/photostream/. Licensed under CC BY 2.0; https://creativecommons.org/licenses/by/2.0/

fitness of the signaler, and perhaps itself. It is a common misconception that communication always consists of a simple signal that is reciprocated with a single response. In fact, communication often uses multimodal sensory combinations of visual, olfactory, tactile, gustatory, electrical (as in electric fish or the duck-billed platypus, *Ornithorhynchus anatinus*), substrate-borne vibrational, and acoustical modes. The use of multimodal signals helps ensure that the message is unmistakable. For example, a cat can swish her tail, pull back her ears, swipe with her claws, and hiss to give an aggressive signal of potential attack, whereas just hissing or swishing her tail is a less clear message.

The focus of this chapter is substrate-borne (vibrational) and acoustic (sound) communication. A signal, for the purposes of this chapter, contains substrate-borne or acoustic information that is broadcast by an individual and is available to be received by another individual. The receiver may be the intended target of the signal or an unintended eavesdropper. Any individual in the environment with the appropriate receptor can receive the signal (Wiley 1983). The receiver of a signal may recognize it as containing information beyond that of just sensing the signal and the presence of the signaler.

### 11.3.1 Communication Concepts

Marler (1961) recognized four functions of signals: identifiers, designators, prescribers, and appraisers. For example, a male seal swims into the territory of another seal and the territory holder sends out a warning call. This call identifies the place and time of the territory holder (identifier), reports that he is the territory holder (designator), warns that the intruder (prescriber) should stop approaching, and allows the intruder to react to his call (appraiser). Smith (1969) expanded this into 12 generalized categories for vertebrates. Since then, with technological and analytical advancements, signal functions have been expanded to include complex displays, either vocal or nonvocal, and the other categories explored below.

Displays are behaviors that use one or several signals. These signals have evolved and become specialized to convey specific information. A classic example of a display behavior is the chest-beating of a mountain gorilla (*Gorilla beringei*), made famous by *King Kong* movies. This signal is given only by the dominant silverback males when he encounters a threat, such as another gorilla male, though the display can be practiced or mocked by the young (Fig. 11.4). The chest-beating forms part of a complex threat

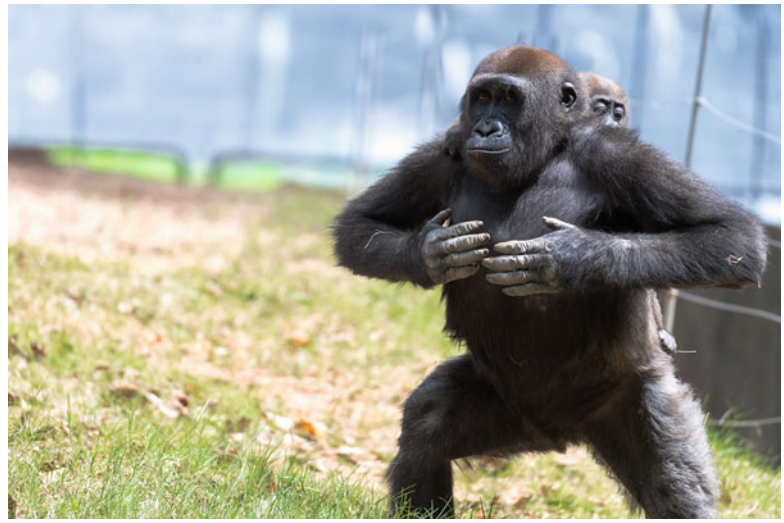


**Fig. 11.4** Displays such as shown by this young gorilla (*Gorilla beringei*) often accompany both vibrational and acoustic communication. Photo "Gorilla Holding Baby Sister and Beating Her Chest" by Eric Kilby; https://www.flickr.com/photos/ekilby/36360289044. Licensed under CC BY-SA 2.0; https://creativecommons.org/licenses/by-sa/2.0/

display, which involves nine steps, and includes both visual and acoustic modalities (Schaller 1964). In other words, the threat display can encompass several different signals.

A similar threatening display is produced by a dog (*Canis lupus familiaris*), drawing back its lips and exposing its teeth (visual), as well as growling (acoustic) (Fig. 11.5). Again, this is a complex display involving multiple steps and multiple modalities. However, displays can be simpler, such as a grasshopper (Orthoptera) scraping its wings as an acoustic signal to indicate location and readiness to mate.

Much of the communication in insects, other invertebrates, and nonmammalian vertebrates such as fish and amphibians, involves stereotyped signals. That is, the signal is produced in a constant form and the response is evoked only by that signal. As a result, this signal/response relationship becomes characteristic of that species. In this way, stereotyped signals can be important in evolution. For example, if a signal influences mate selection, then a slight alteration in the signal could lead to failure to reproduce, or if mating is successful, it might give rise to a new species.



**Fig. 11.5** Yellow Labrador retriever growls at a border collie, while using a mix of visual displays and vocalizations; the collie responds. "Growl" by smerikal is licensed under CC BY-SA 2.0; https://commons.wikimedia.org/wiki/File:Labrador_Growl.jpg

### 11.3.2 Biotremology

Vibrational behavior in animals has gained momentum in general awareness and research in the last few decades (Narins 1990; Hill 2008; Cocroft et al. 2014; Hill et al. 2019). Any sort of motion of a living organism produces vibrations in the various media around them, including the soil, air, plants, water surface, or spider webs. Some vibrations can be signals, while others are incidental cues not produced purposefully, or to benefit the sender. The rather new branch of behavioral biology studying vibrational communication is called "biotremology" and is concerned with substrate-borne mechanical waves used as a communication channel (Hill and Wessel 2016). In contrast to airborne sound, which consists of pressure waves only (see Chap. 4, Sect. 4.2.2), in solid substrates mechanical energy can travel in several waveforms, especially at the surface (i.e., the boundary between two distinct media; Fig. 11.6). Surface-borne waves are of special interest as most animals that make use of vibrational communication receive the signals by detector organs. These organs are in contact with a substrate surface, be it the ground, the surface of plant stems and leaves, or the water surface.

In addition to pressure waves (P-waves) and shear waves (S-waves) traveling inside the body of a solid (see Chap. 4), we have at the substrate surface Rayleigh waves (R-waves) and Love waves (L waves). Both R- and L-waves show particle oscillation perpendicular to the direction of the wave, but different propagation characteristics. P- and L-waves, for example, both have a higher propagation velocity than R-waves. Animals who can detect those waveforms differently could localize the source of these waves—be it a communication partner, a predator, or prey.

In 1979, Brownell and Farley showed that scorpions localize their prey by using differences in the propagation velocity of P- and R-waves (150 m/s:50 m/s), which they perceive using

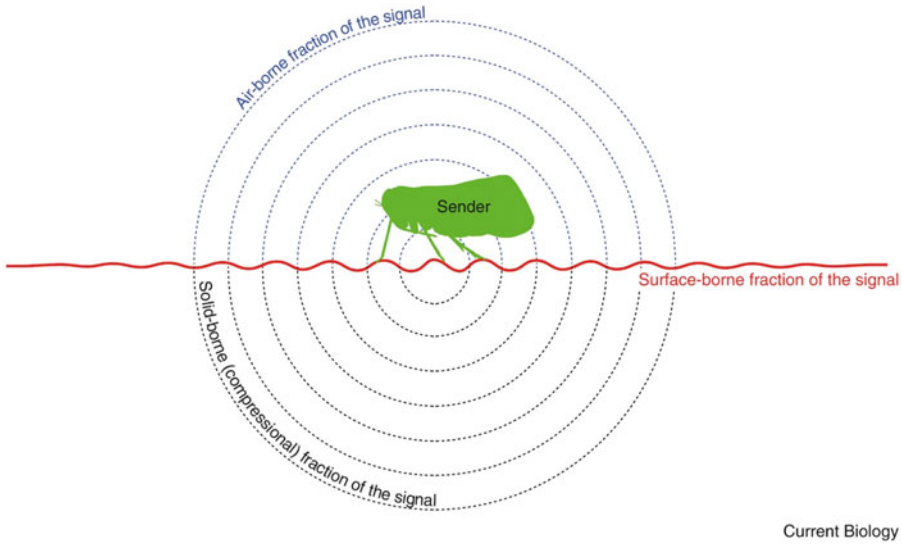



**Fig. 11.6** Mechanical wave forms produced by a signaling plant-dwelling insect. A planthopper is one of the small relatives of the cicadas. It has a tymbal organ to produce vibrations, which are transferred through its legs, then the thin air layer between its body and the plant surface, to the plant on which it is sucking fluids. By doing this, the planthoppers produce a very faint sound, which can be propagated through the air or soil. The planthopper tymbal organ is homologous to the "drumming organ" of the large singing cicadas. Tens of thousands of these smaller hemipteran bugs use tymbal organs to produce "silent songs." Reprinted by permission from Elsevier. Hill P SM, Wessel A (2016). Biotremology. Current Biology 26, R181–R191; https://doi.org/10.1016/j.cub.2016.01.054 © Elsevier, 2016. All rights reserved

different sensory organs (tarsal hair receptors v. basitarsal slit sensilla). That was a significant discovery on the path to biotremology. Until then, the substrate the scorpions use, loose sand, was considered as not fitting for the transmission of vibrational signals, nor for the differential detection of different waveforms. Since the establishment of the view that a host of natural substrates are suitable for vibrational communication, a great number of (apparently) well-known behaviors are now seen in a new perspective, and new discoveries are made for almost all animal groups with increasing frequency (Hill et al. 2022).

The production of vibrational signals nor cues can be accomplished through different forms: drumming (any sort of percussion event where a body part impacts the substrate of soil or a plant or water, etc.), tremulation (a body shaking/trembling that does not strike the substrate as the signal travels through the signaler's legs to the surface on which they are standing), stridulation (rubbing together a specialized file and scraper, which may be found on a variety of body parts), buckling of tymbal organs in animals that have them, vocalizations and perhaps others, such as scraping a surface while signaling, or even scratching against a tree, or rolling on the ground. Some of these signal production mechanisms, such as drumming, stridulation, and vocalization, always produce both a substrate-borne (vibrational) and an airborne (acoustic) component with a single action, even if only one of the potential signals is capable of eliciting a response in a receiver.

Arthropods, and especially insects, show the greatest variety of specialized organs to produce vibrational signals. All mentioned means of vibration production, except for vocalization, are present in several groups of arthropods and may have evolved several times, independently. For a subgroup of the insect order Hemiptera, the Tymbalia or tymbal bugs, comprising tens of thousands of species including plant- and leafhoppers, cicadas, and true bugs (Heteroptera), vibrational communication is known to be evolutionarily old and ubiquitous (Hoch et al. 2006; Wessel et al. 2014).

In mammals, most vibrational signals are produced by drumming or vocalization. Curiously, the vibrational communication of the largest land animal, the African savanna elephant (*Loxodonta africana*), was discovered by O'Connell-Rodwell in the 1990s, when she noticed peculiar behaviors. A freezing behavior in the elephant and change in orientation, without an apparent cause, nevertheless reminded her of the behaviors of the tiny planthoppers whose vibrational communication she had studied earlier (Fig. 11.7). O'Connell-Rodwell and colleagues demonstrated that the signals the elephants generate with low frequency "rumbles" (about 20 Hz) could be very useful for intraspecific long-distance communication (O'Connell-Rodwell et al. 1997, 2000).

Also, drumming is a type of long-range vibrational signal production. For instance, drumming by prairie chickens (*Tympanuchus cupido*) can be detected up to 5 km away from the source (Jackson and DeArment 1963). Kangaroo rats (*Dipodomys deserti*, *D. ingens*, and *D. spectabilis*) drum the soil surface (seismic communication) with their feet to communicate such things as territorial ownership, their competitiveness, and their presence and location to other kangaroo rats (Fig. 11.8, Randall 1984; Randall and Lewis 1997; Cooper and Randall 2007).

Many species of marsupial kangaroos (Macropodidae) are known to produce a foot thump when confronted by predators. The intended recipient of the vibration is not known and could be either a predator or other kangaroos (Narins et al. 2009). Sheep and many other ungulates stamp their feet when frightened or aroused in other ways.

As every movement of an animal cause particles in the surrounding media to oscillate and evokes all possible sorts of mechanical waves, it is the mechanism of reception of mechanical signals or cues that defines acoustic vs vibrational communication. It also follows that every act of communication establishes—at least potentially—a complex communicational network in the realm of the "acousto-vibro-active-space," whereby the active space for vibrational signals can be surprisingly wide, even bridging air gaps (Fig. 11.9; Virant-Doberlet et al. 2014;

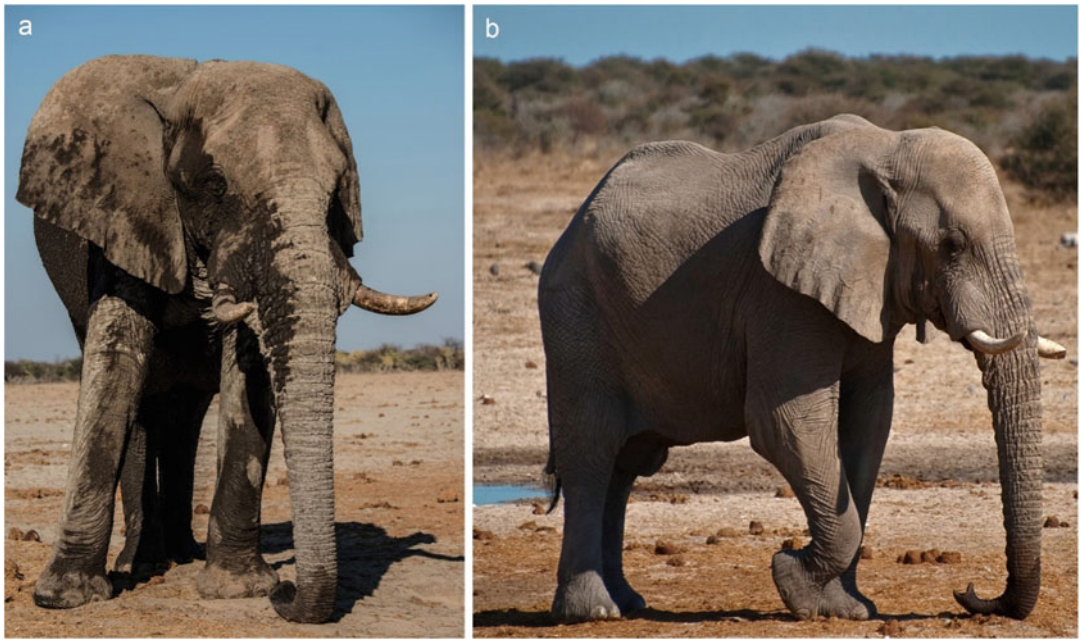



**Fig. 11.7** Elephant vibration detection posture. (**a**) To detect a signal, an elephant appears to focus solely on somatosensory detection via receptors in the trunk. Its ears are relaxed suggesting no airborne assessment for signals. (**b**) Elephant vibration detection posture, where it appears to be using its toenails and trunk to assess a ground-borne signal. Again, its ears are not fully extended. This suggests it uses both bone conduction through the toenails and a somatosensory pathway through Pacinian corpuscles in the trunk for signal detection. Elephants may also lean forward on their front legs with ears flat, sometimes lifting one of the front feet off the ground (possibly for triangulation or better coupling). If focused on an acoustic signal, an elephant will hold its ears out and scan its head back and forth in the general direction of the sound. Reprinted by permission from Springer Nature. Biotremology: Studying vibrational behavior, edited by P. S. M Hill, R. Lakes-Harlan, V. Mazzoni. P. M. Narins, M. Virant-Doberlet and A. Wessel, pp. 259–276, Vibrational communication in Elephants: A case for bone conduction, C. O'Connell-Rodwell, X. Guan and S. Puria; https://link.springer.com/chapter/10.1007/978-3-030-22293-2_13. © Springer Nature, 2019. All rights reserved



**Fig. 11.8** Kangaroo rats (genus *Dipodomys*) produce seismic signals by drumming the soil surface with their large hind feet. (left) Photo of "Kangaroo Rat by Stuart Wilson" by cameraclub231 is licensed under CC BY 2.0 (https://www.flickr.com/photos/135081788@N03/49936422922). (right) Ord's Kangaroo rat (*Dipodomys ordii*). Photo of "Two Ord's Kangaroo rats, Alberta" by Andy Teucher licensed under CC BY-NC 2.0; https://www.flickr.com/photos/63265212@N03/8736679123

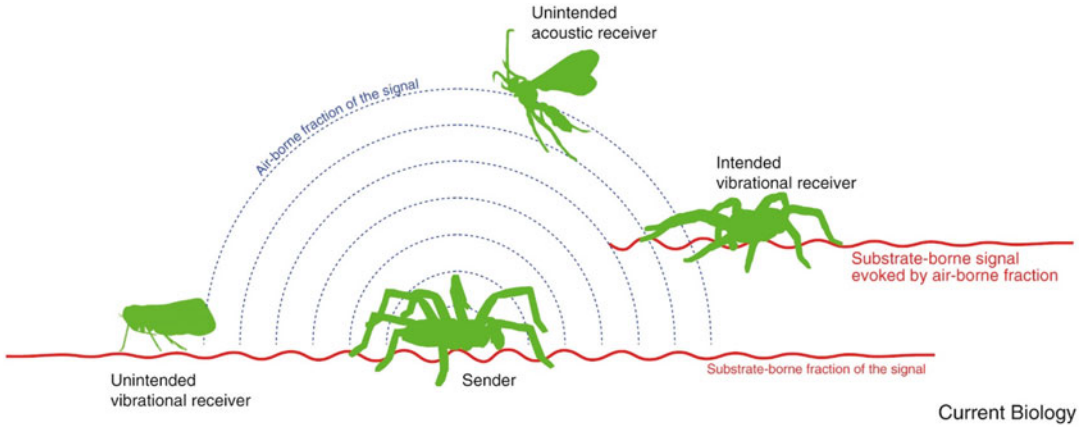



**Fig. 11.9** Types of communication acts by a vibrational signaler. The signaling lycosid wolf spider establishes vibrational communication with a conspecific receiver, even one that is not on the same substrate as the sender. Likewise, a vibrational communicating prey (e.g., a planthopper) and an acoustically orienting parasite (e.g., a braconid wasp) are eavesdropping on the spider whereby establishing a complex communication network. Reprinted by permission from Elsevier. Hill P SM, Wessel A (2016). Biotremology. Current Biology 26, R181–R191; https://doi.org/10.1016/j.cub.2016.01.054. © Elsevier, 2016. All rights reserved

Mazzoni et al. 2014; Gordon et al. 2019). On an ecosystems level, we have begun to think of, and to study, a whole complex multilevel vibroscape (Šturm et al. 2021).

Despite the importance of reception mechanisms for the study of vibrational communication, they are, for now, the least understood aspect in biotremology. Arthropods have in their bauplan—in every body segment and at every joint of their legs—mechanosensitive stretch organs (chordotonal organs) that are responsible for body and movement control, but could also pick up environmental vibrations. In some groups, such as grasshoppers, crickets, and cicadas, chordotonal organs have evolved into ears with a tympanum attached to one end of the stretch organ. It is hypothesized that in every such case these hearing organs transformed through an evolutionary intermediate stage of vibration receptors, i.e., vibrational reception is evolutionarily older than hearing.

A recent breakthrough was the demonstration of the complete pathway, from signaling through reception, to perception, and response behavior, of the vibrational component of the courtship of the fruit fly *Drosophila melanogaster*. It is the vibrational signaling of the male that triggers the female to freeze at the end of the courtship, facilitating copulation (McKelvey et al. 2021). The male's vibrational signals are transmitted through the common courtship floor—overripe fruits—and were picked up by a subset of neurons of the female's femoral chordotonal organ. By genetic knockout experiments of several mechanotransducer ion channels, McKelvey (et al.) also identified a protein involved known to be responsible for gentle touch sensitivity in vertebrates—suggesting a deep evolutionary origin of vibrational communication.

In several cases, we need to consider a bimodal acousto-vibrational communication on the signal production as well as on the reception side that results in a complex perception of the environment outside of the experience of human beings. Elephants, for example, produce low-frequency signals by vocal "rumbles" and "foot stomps" that produce airborne vibrations (sound) as well as seismic waves (O'Connell-Rodwell et al. 2000). New findings point to a simultaneous monitoring of the signaling by three reception pathways: sound hearing by the ear's tympanum, bone conduction hearing, and somatosensory detection via receptors in the trunk (Fig. 11.7; O'Connell-Rodwell et al. 2019). In this way, the

overall chance of detecting a signal at all in a heterogeneous environment is improved, and the animals could also make use of the different propagation velocities for assessing the distance to the source of the signal.

### 11.3.3 Diversity in Communication

Recent evidence indicates that many messages may be conveyed auditorily in nonhuman primates when the larynx is not used. These commonly take the form of rumbling of the stomach, farting, breaking sticks, swishing of grass, sounds during digging or flying, and others. In fact, many sounds made by an individual can carry information to those who hear, but the question is whether they are used for communication. These sounds could just be the result of physiological or environmental adjustments that the sender may or may not be able to control, or that are not recognized as significant in communication. One example is surface behavior in humpback whales. Humpback whales can launch their body out of the water, turn, and splash down on their side or back (breach), slap the water with their pectoral fins, tail flukes, and even their head. These produce loud "bang" sounds, thought to be used as communication signals during periods of high underwater noise when vocal signals are not as effective (Dunlop et al. 2010).

In general, the use of these sounds for communication has not been given much research time to date, except for cases where they have been ritualized to carry information to others. For example, we do know, from centuries of hunter's anecdotal evidence, that a hunted antelope, elephant, or even a rhino, will move much more carefully to not make a sound when it is being hunted, compared to when traveling/grazing in a group (e.g., Baze 1950). If this is the case, the individual must recognize that the sound will carry a message (Heyes and Dickinson 1990).

In invertebrates and non-primate vertebrate animals, ascertaining whether or not these signals are being used for communication is more of a challenge. Each movement of an animal's body creates vibrations that propagate through the environment, and production of these vibrations cannot be eliminated by the individual, even if walking more softly does lower the amplitude. Therefore, we can be certain that in both vertebrate and invertebrate predators, a substrate-borne vibration or sound that alerts potential prey of the presence and direction of movement of the predator is not communication. In animal communication, we refer to this class of unintended information as a cue. On the other hand, we may also be familiar with a hunting dog moving through a meadow and flushing birds on the ground into flight with the result that the hunter can shoot them. We simply do not know if this sort of behavior exists in a more natural less domesticated setting.

## 11.4 The Advantages and Disadvantages of Vibrational and Acoustic Communication

Substrate-borne vibrational and acoustic signals are used in communication by almost all invertebrates and vertebrates. Sometimes each type of signal is used by a single species but in different contexts. There are many examples of the two being used across animal taxa in the same basic context. Some major groups of animals have evolved a heavier dependence on one than the other. For example, only as recently as 2015 did we observe the first described substrate-borne signaling in mating birds (Ota and Soma 2022) and in the very well-studied fruit fly *Drosophila melanogaster* (McKelvey et al. 2021), both of which were well-known for acoustic and visual signaling. These signals are essential for many species to find a mate, keep in contact (such as between mother and young), maintain territory, warn conspecifics of predators, link food location, reinforce social living, communicate emotional state, and many other types of information (Bradbury and Vehrencamp 1998). For any animal, being out in the world advertising your presence has many advantages, but it also has its disadvantages. The advantages of using vibrational and acoustic communication signals are

essentially the same. There is no need for light—so signals can be detected at night. Sound can flow around obstacles, so acoustic signals can be heard anywhere and anytime, and even though the substrate filters vibrational signals and cues in ways that are difficult to predict, they still can be detected without respect to time. Compared with other signals, most vibrational and acoustic signals do not need a great deal of energy to produce. Because of the physics of signal propagation, vibrational and acoustic signals can travel over long distances. For instance, in primates, the roaring of howler monkeys (genus *Alouatta*) can travel up to 1 km.

However, there are disadvantages to vibrational and acoustic communication. These include energetic and developmental costs, such as requiring special structures for signal production and reception. Being able to produce a loud signal often requires new, and possibly elaborate structures, such as the larynx of vertebrates and the melon of sperm whales, *Physeter macrocephalus*). Invertebrates have also evolved specialized structures, such as the stridulatory apparatus in insects, which requires a receptor such as the subgenual organ (for substrate-borne vibrations) and the ear (for sound) to pick up the messages. Many animals have evolved specialized receptors to detect substrate-borne vibration signals (Pacinian corpuscles, Meissner's corpuscles, Eimer's organ; Narins and Lewis 1984; Narins et al. 2009).

The disadvantages of signaling can, however, be subtle—such as a wasted broadcast when there is no one to receive it or alerting others and then being overcome by a predator. "Blurting out" who and where one is means others can find you. By listening in, these others, or unintended receivers, which could be predators, prey, or even eavesdropping conspecifics, can obtain valuable information about the signaler. This may come at a cost to the signaler. If the unintended receiver is a predator, the cost is obvious: by listening in on the sound signals, the predator can recognize the signaler as prey and locate it. Conversely, prey can be alerted to, and identify, a signaling predator and its location, thus making it easier for prey to avoid predation. A conspecific eavesdropper can gain important information about the signaler/receiver relationship without having to directly take part in the interaction. Siamese fighting fish (*Betta splendens*), for example, eavesdrop on fighting males to gain information about their strength, which they then use in future interactions (Oliveira et al. 1998; Peake and McGregor 2004). To add further complexity, the presence of an eavesdropper audience can affect communicative interactions and force signalers to change their signaling behavior according to who else may be listening in. This is known as the audience effect and was first documented in a study of domestic chickens (*Gallus gallus;* Evans and Marler 1991, 1994).

Despite these and other disadvantages, it is obvious that substrate-borne vibrational and acoustic communication and all that they entail have provided extraordinary benefits in competing, surviving, and propagating the next generation. The stories of the development of vibrational and acoustic communication are ongoing and much knowledge about the mechanisms, meanings, and extent of these systems is yet to be discovered.

## 11.5 The Influence of the Environment on Acoustic and Vibrational Communication

For the most part, animals do not sit in a studio, acoustic lab, or anechoic chamber when signaling acoustically or with substrate-borne vibrations. They are usually in a natural environment subject to atmospheric and other conditions. Signals may be affected by spatial separation, movement of the caller, and they may even vary spatially or geographically. Environmental noise is a significant factor influencing animal signaling behavior. While few studies to date have addressed vibrational environmental noise, this topic is the focus of a recent review of both terrestrial and marine anthropogenic noise topics and literature, including previously unpublished case studies that can

be used as guides for future work (Roberts and Howard 2022).

### 11.5.1 Atmospheric Conditions

Atmospheric conditions, which include changes in temperature and wind, exert powerful and predictable influences on animal sounds. These influences can cause the ability to detect a signal to change rapidly. The transmitting of a signal may be prolonged or modulated by topography, regional weather, seasonality, and climate. Mammalian carnivores, such as coyotes (*Canis latrans*) and wolves (*Canis lupus*), live in areas with nocturnal lower temperatures (David Mech and Boitani 2003). These animals show crepuscular calling to maximize their chances of being heard over the longest possible distances. Vibrations in the soil or other substrates due to wind or rain can also interfere with normal signal production and reception to the extent that individuals will stop courtship displays under windy or rainy conditions.

### 11.5.2 Masking Sounds

Masking sounds are environmental sounds, such as a stream, wind moving through the trees, and sounds from other animals, which cover, or dilute, the signal. In birds and other animals, spatially separating a signal from a masking sound is one way to improve signal detectability. If the signal and masking sound are separated spatially, the receiver can focus efforts to hear the signal. This "spatial release from masking" has been demonstrated in the behavior and physiology of the northern leopard frog (*Lithobates pipiens*) (Ratnam and Feng 1998). Bee (2007) showed that female Cope's gray treefrogs (*Dryophytes chrysoscelis*) approached a target signal more readily when they were spatially separated by 90° from a masking sound, implying this spatial separation aided with signal reception. Spatial release from masking has also been shown to occur in budgerigars (*Melopsittacus undulatus*; Dent et al. 1997) and killer whales (*Orcinus orca*; Bain and Dahlheimm 1994).

A similar mechanism to spatial release from masking is known as the cocktail party effect. Here, the receiver focuses its attention on the signaler, while selectively filtering out other stimuli such as other sounds. At a party, humans can "tune in" to one conversation when many are taking place. Many frogs and songbirds have also been shown to successfully communicate in noisy party-like situations. Frogs can recognize, localize, and respond to signals within a cacophony of chorusing (Gerhardt and Bee 2006; Wells and Schwartz 2006). Songbirds are able to recognize conspecific song and songs from other species within a dawn chorus (Benney and Braaten 2000; Hulse et al. 1997). Reunited offspring and parents within a noisy colony clearly occur successfully in penguin colonies (Aubin and Jouventin 1998).

The above mechanisms demonstrate how the receiver overcomes masking sounds to improve signal detectability. Another way to improve signal detectability is for a signaler to change the way it calls. For example, a signaler could increase its call amplitude, call duration, and/or call at a different frequency. These changes are collectively known as the "Lombard Effect." The Lombard effect has been demonstrated in species such as the Japanese quail (*Coturnix japonica*; Potash 1972), budgerigars (Manabe et al. 1998), chickens (*Gallus gallus domesticus*; Brumm et al. 2009), nightingales (*Luscinia megarhynchos*; Brumm and Todt 2002), white-rumped munia (*Lonchura striata*; Brumm and Zollinger 2011), and zebra finches (*Taeniopygia guttata*; Cynx et al. 1998) and even in large whales such as the humpback whale (Dunlop et al. 2014).

### 11.5.3 Geographic Variation and Dialects

Changes in the environment may lead to geographic variation, and this variation can eventually separate animals within a species into different populations. It should be noted that geographic variation is not necessarily due to changes in the environment. While this is occurring, geographic separation can lead to the

formation of dialects. A dialect can evolve where species dispersal is occurring and their acoustic contact with each other becomes limited (Slater 1986, 1989). As a result, individuals within a species population may exhibit similar sounds to each other, but these sounds may be quite different in structure to other separated and more distant populations (Catchpole and Slater 2008; Gannon and Lawlor 1989). This results in within-species vocal variation.

Dialects are also known from biotremology studies. For example, the well-known southern green stink bug (*Nezara viridula*) has spread throughout the world (except for the Arctic and Antarctic) from its native Ethiopia in the past 100 years. Geographically isolated populations (e.g., California and Florida in the United States, the French Antilles, Australia, Japan, Slovenia, and France) have distinct differences in duration and repetition time of male and female signals. Individuals appear to be able to recognize adults from other populations but prefer to mate with those of their own dialect/population (Virant-Doberlet and Čokl 2004).

The study of population dialects offers a means to explore the causes and the functions of signal variation and change (Henry et al. 2015). Geographic variation in acoustic signals can reflect historical evolutionary changes within species. Not only can these signals be used to assess links between geographic variations and population connectivity, but they can be used to provide important information for the conservation of a species. For example, geographic variation in calls could indicate how birds disperse through a fragmented habitat, meaning the study of dialects can be used as a noninvasive tool to assess population connectivity (Kroodsma and Miller 1982; Amos et al. 2014).

The formation of dialects can occur through several mechanisms; as a result of a side-effect or "epiphenomenon" of learning via incorporating copying errors (such as adding or omitting parts of the call), due to structural changes to call elements through drift, or as a possible indicator of the level of behavioral or genetic variation in a population (Baptista and Gaunt 1997; Catchpole and Slater 2008; Podos and Warren 2007; Keighley et al. 2017). Another mechanism that helps maintain variable acoustic dialects is social adaptation. Social adaptation refers to the ability to adjust behavior to a prevailing pattern in a population. Migrating birds, for example, learn calls quickly (Salinas-Melgoza and Wright 2012), which provides reproductive benefits due to acoustic familiarity by potential mates (Catchpole and Slater 2008; Farabaugh and Dooling 1996). In this way, newly arriving immigrants fit in quickly and do not insert changes to bird songs of the residents, thereby maintaining the local dialect.

Vocal dialects can act as precursors to genetic isolation (e.g., in coastal US chipmunks, genus *Neotamias*). Dialects can also be maintained over time if the populations are separated and have little acoustic contact. This separation can be reinforced by geographic boundaries, or other isolation mechanisms, that reduce breeding chances (Gannon and Lawlor 1989). Examples include the pika (*Ochotona*), grasshopper mice (*Onychomys*), white-crowned sparrows (*Zonotrichia*), prairie dogs (*Cynomys*), and bats (*Myotis evotis*), which have all been shown to exhibit dialects due to geographic variation. Several species of birds, such as the chaffinch (*Fringilla coelebs*), have been identified as having song dialects and therefore are described as having distinct "cultures" (Slater 1981). One of the most striking examples of cultural influences is the rapid spread of new humpback whale songs across the South Pacific basin. All male humpback whales within a population generally conform to the same song pattern, making it a cultural trait. These song types move eastward across the South Pacific basin in a series of cultural waves at a geographic scale unparalleled in the animal kingdom (Garland et al. 2011).

Behavioral repertoires are malleable—that is, they are affected by the environment, learning, and interactions within a population. Variants in signal characteristics are no exception (Brumm et al. 2009). Thus, signal characteristics can act as precursors to variants in other genetic characteristics, and eventually, speciation.



**Fig. 11.10** Hoary bat (*Lasiurus cinereus*). "Hoary bat" (https://www.flickr.com/photos/33247428@N08/48546621027) by Oregon State University is licensed under CC BY-SA 2.0; https://creativecommons.org/licenses/by-sa/2.0/

Notably, O'Farrell et al. (2000) examined nearly 2500 calls from 43 sites in Hawaii and mainland United States for the Hoary bat (*Lasiurus cinereus*; Fig. 11.10). They found some geographic variation within the calls, but the variation could not be explained by isolation (mainland distance of about 2300 miles (3800 km) from the proximity of San Francisco, CA, USA and Honolulu, Oahu, Hawaii, USA). They were unable to exclude the effects of context, behavior, or in some cases low sample size. Bats of this species, regardless of where they were recorded, could be identified as *L. cinereus*. In other words, these bats were showing variations in call structure and behavior but had not yet evolved into different species.

There are instances in which different species have evolved. Several studies in mammals have found that research into the geographic variation of acoustic signals is important taxonomically by discovering cryptic species. Chipmunks (*Neotamias*) occurring mostly along the US coasts of California, Oregon, and Washington were thought to be one species (*Eutamias townsendii*) with several subspecies. The species was characterized mostly by cranial and pelage features. It was not until localities throughout the range of the four subspecies within *E. townsendii* were sampled acoustically, and examined statistically, that variation of the calls was shown to be dramatic enough to warrant elevation to four distinct species. Originally based on acoustic data, this was confirmed by genitalia and genetic information (Gannon and Lawlor 1989; Sutton and Nadler 1974; Sullivan et al. 2014).

## 11.6 Information Content or the Meaning of Signals

Vocal signals can be used to provide (a) static information about the species, including the size and shape of the vocal apparatus, or (b) dynamic information, that is, the motivational state of the sender. Vocal signals can be context-dependent, where the same call can mean different things in different situations, or context-independent, where the call has a specific meaning whatever the context. Species recognize one other from their vocalizations, and produce signals related to various situations such as alarm calls in the presence of a predator, distress calls when separated from a parent, singing and chorusing to attract or deter conspecifics, or reflect behavioral changes. The question then arises; how does the recipient know what the caller means in that situation? The answer is, at least in birds and mammals, the receiver assesses call meaning by observing the sender and the context in which the signal is sent.

### 11.6.1 Static Information

In addition, the anatomy of the vocal apparatus in mammals determines features of its sounds, and these features correlate with the animal's body size (Fitch 1997 in rhesus macaques, *Macaca mulatta*). Larger lungs can produce longer vocalizations. Vocal folds that are longer and thicker produce sounds at lower fundamental frequencies (for example, pika, *Ochotona alpina*; Volodin et al. 2018). The longer vocal tract concentrates the energy in the lower frequencies (Ey et al. 2007). Thus, correlations have been found between an animal's vocal tract length,

body mass, and formant dispersion (e.g., domestic dog, *Canis lupus familiaris*, Riede and Fitch 1999; southern elephant seals, *Mirounga leonina*, Sanvito et al. 2007).

As a result, information about the sender's body size, sex, age, and sometimes rank can be acquired from their vocalizations. Sounds from small or young animals are typically higher in frequency than those of larger or older animals (see Riondato et al. 2021 for an exception). Sometimes rank information is used by females selecting males. For example, the "roar" of the male Red deer (*Cervus elaphus)* contains information on its sex and size. The larger the animal, the lower the frequency of the roar. Females chose mates based on their roar and have been found to prefer the roars of larger males (Charlton et al. 2007). The signaler's dominance rank can also be signaled using size-related formants (e.g., male fallow deer, *Dama dama*, Vannoni and McElligott 2008; and baboons, *Papio ursinus*, Fischer et al. 2004). As the sender's features do not change (e.g., their sex), or change slowly over time (e.g., their size or age), it is known as static information.

### 11.6.2 Dynamic Information

A second type of information is known as dynamic. This information relates to the sender's motivation or arousal. Dynamic, or context-dependent calls, follow a motivational code (Morton 1977). A loud or long sound, for example, is associated with the signaler experiencing high arousal that may be due to aggression, fear, frustration, distress, or pain. Signalers in hostile contexts tend to emit longer, lower-frequency "harsh" (broadband) sounds which can signify signaler size. These sounds function to mediate aggressive interactions between it and the receiver. High tonal sounds, that mimic infant sounds, are more likely to be emitted in appeasing (fearful) contexts given they potentially have an "appeasing" effect on the receiver. Distress calls (often "scream" or "whistle-like" vocalizations) are used when "fear" and "aggression" are conflicting motivations. A short quiet signal is often associated with pleasure, close contact between animals that like each other (such as mother to young), or between social partners when close (Morton 1977).

Affiliative calls can indicate a welcoming, or "I am fond of you" context. For example, familiar elephants meeting each other after a long separation may trumpet for pleasure/joy (a high state of arousal). They also murmur to a friend, infant, or person they like who has been close, indicating a low level of arousal but a similar emotion (Kiley-Worthington 2017).

Aggressive calls include territorial calls and calls used as threats, and like affiliative calls, the agnostic call structure can change because of arousal. A highly aroused bull (*Bos taurus)*, for example, will give visual signals: pawing, lowering his head withdrawing his chin and rubbing his horns in the earth, at the same time as roaring. At the highest level of threat, the roar has a vocalized inspiration as well as a vocal expiration known as a "see saw" call (Kiley 1972).

### 11.6.3 Context-Dependent Meanings

Context-dependent communication is where the same signal may be used in different contexts but has different meanings. For example, a male eastern kingbird (*Tyrannus tyrannus*) emits a "kitter" call-in three different contexts: (1) when the bird is indecisive or concerned about attempting to approach some object (to perch, mate, or toward another bird), (2) when lone males fly from perch to perch in a new delimited territory, or (3) as an appeasement signal by the male when approaching his mate. Another example is the familiar roar of a lion (*Panthera leo*) that—from the viewpoint of a human—is a spectacular vocal display during aggressive interactions. However, the call also helps individuals belonging to the same pride find, and identify, each other and can serve as a bonding signal for members of a pride to gather. It can also separate neighboring pride.

Affiliative calls can also be food calls (Kondo and Watanabe 2009). Food calls can be context-dependent given these signals are directed at other conspecifics and can indicate the presence of

food. The variation in these food calls can indicate food a quality and quantity. For example, spider monkeys (genus *Ateles*) are known to produce a higher call rate in response to greater quantities and quality of food. Acoustic signals can attract group members to food locations and these calls can also be used to protect the food resource from others (Clay et al. 2012). These authors examined food-associated calls made by some birds and mammals (see page 326, Table 11.1 in Clay et al. 2012) and found that most species did not produce unique calls for different foods. More commonly, signalers varied their calling rate to advertise food quality or abundance.

Therefore, context-dependent vocalizations may not necessarily convey information about the type of situation but can act as an analogue system to inform the recipient about the general level of arousal of the sender, and consequently, how (or if) to respond. In some species, calls are graded, meaning that there are intermediates between one call and another. Humpback whales, for example, use a repertoire of graded signals and the use of these signals is likely related to the motivation and arousal of the signaler (Dunlop 2017). "Grumbles" and "snorts" are used by females and their calf while migrating by themselves and presumably in a low-arousal context. Female–calf pairs can be joined by male escorts and form a competitive group, where males are fighting for access to a breeding female. In these groups, where arousal level is much higher, "grumbles" turn into harsh sounding "roars" and "purrs," and become more modulated to sound more like "groans" and "moans."

Different levels of graded calls can be given in one situation. For example, cattle may give a low "mmmmm" call when in close contact with other cattle. On opening its mouth, the sound has an added syllable: "en" to "mmen." When it is sufficiently aroused, a "hh" syllable is added, which is the result of letting the remaining air out of her respiratory track. This can change even further with higher excitement or arousal by being repeated. Finally, at the highest level of arousal, the inspiratory phase of the call is also vocalized (Table 11.1). This is a very different type of auditory communication from context-independent calls such as human language where auditory communication can reflect either or both and environmental contexts or come from some thought or idea generated by cognition.

### 11.6.4 Species Recognition

To be sure that the call maintains the same structure (and can therefore be recognized as having the same message), there are a number of measures including call interval, maximum frequency, minimum frequency, fundamental or

**Table 11.1** The variety of situations that give rise to the major call types of *Bos taurus* (reproduced from Kiley 1969)

| Situation/call | mm | men | menh | (m)enENh | SeeSaw A (no inspir) | SeeSaw B (+inspir) |
|---|---|---|---|---|---|---|
| Confident greeting | + | + | + | | | |
| Greeting equals | + | + | + | + | | |
| Defensive threat | | | | + | + | + |
| Aggressive threat | | | + | + | + | + |
| Fear | | | | + | + | |
| Close contact retain | + | | | | | |
| Tactile stimulation | + | | | | | |
| Isolation | | + | + | + | + | + |
| Startle | | | | | | |
| Pain/fear | | + | + | + | | |
| Frustration | + | + | + | + | + | + |
| Anticipation pleasant | + | + | + | + | + | + |
| Anticipation unpleasant | + | + | + | + | + | + |
| Disturbance | + | + | + | + | + | + |

predominate frequency, call length, duration, amplitude or loudness, and the repetition rate found in both acoustic and vibrational signals. These characteristics, combined with the presence of harmonics, form patterns that are often characteristic of a species or individual. As a result, other animals are likely to be able to identify individuals from their calls, as we can with human voices. For example, many species of vespertilionid bat can be identified by time and frequency characters measured from their echolocation calls (Gannon et al. 2003). Individual recognition is also evident in bats. Playback responses in common vampire bats (*Desmodus rotundus*) suggested they vocally recognized individual bats, given they were biased toward callers that had fed them more (food sharing), but not biased toward kin (Carter and Wilkinson 2016). Crickets (*Teleogryllus* spp.) can be differentiated based on the amplitude and repetition of their call, not just their call "note" (that is, the fundamental). The mean frequency of this signal is approximately 4 kHz, but the pattern and call rate increase as the cricket's motivation changes from "calling" to "encountering" to "fighting" to "courtship" and finally "copulating."

### 11.6.5 Context-Independent Meanings

Some calls in animals, like human language, have a specific meaning, whatever the context. These calls often include alarm calls used to alert a group to danger of an approaching predator, territorial invader, or other "alarm" in the caller's environment. The alarm call may elicit a response by recipients to retreat, freeze in place, or conduct defensive behavior. Slobodchikoff et al. (2009) discussed the complexity of alarm calls in prairie dogs (*Cynomys gunnisoni*) in the southwestern United States. He and his students have found that prairie dogs are precise in their signaling and can communicate a description of the predator, its size, its speed, and even its color. Wild boars (*Sus scrofa*) use context-dependent calls, such as "grunts" and "screams," whose meanings relate to the context, also emit a specific "warning bark"—a context-independent short sharp call that is difficult to locate as an alarm call (Kiley 1972). This alarm call works to conceal the position of the signaler but conveys that a disturbing object has been sighted.

The importance of altruism (or lack of it) when vocalizing has been investigated within the context of emitting alarm calls and food calls. For example, studies have shown that, even those calls that are difficult to locate (ventriloquial calls), will increase the chances of being detected by a predator (Fig. 11.11). However, studies on kinship and altruism have yet to relate the ease of locating an alarm call by a predator to the rate of vocalizations and to actual predation (Reznikova 2019). Still, it seems that coterie members of prairie dogs (*Cynomys ludovicianus*) alert others to the presence of potential predators using alarm calls, and that these alarms significantly reduce predation (Wilson-Henjum et al. 2019).

Functionally referential signals are those that provide very specific information. They are structurally distinct and reflect a stimulus-specific meaning used only in a very specific set of circumstances. Most alarm calls are nonspecific, but the vervet monkey (*Chlorocebus pygerythrus*), uses a lexicon of four or five sounds to identify the type of intruder. When a major bird or mammal predator is nearby, the vervet produces a "chirp" and "bark" (Strusaker 1966). When a snake is nearby it evokes a special "chutter" call, a minor bird or mammalian predator is indicated by an abrupt "uh" or "nyow" sounding signal, and a major bird predator elicits a "rraup."

Distress calls can be context independent, such as the calls used by young to attract adults to their location. African wild dog (*Lycaon pictus*) pups, for example, emit a "lamenting call" when they are deserted by their parents. Precocial birds, such as domestic fowl, ducks, or geese, "pipe" in the same way as when they are cold or hungry. Young, collared lemmings (*Dicrostonyx groenlandicus*) emit ultrasonic chirps when they are abandoned, cold, or feel as if they are in danger (Sales and Pye 1974). Young primates,



**Fig. 11.11** Young prairie dogs (*Cynomys ludovicianus*) at Rocky Mountain Arsenal National Wildlife Refuge, Commerce, CO, USA. One pup giving a yipping call. US Fish and Wildlife Service Photo Credit: Rich Keen at RMA; https://commons.wikimedia.org/wiki/File:Yipping_Prairie_Dog_Pups.jpg. Licensed under CC BY 2.0; https://creativecommons.org/licenses/by/2.0/

including humans, shriek or scream when threatened or abandoned.

### 11.6.6 Songs

Songs are composed of call notes that have been elaborated in structure and length. The main function of song is to identify the singer as a member of a species, sexually mature, on a territory, prone to territorial defense, and ready for courtship. Song refers to the melodic quality (with harmonics) of songs, as opposed to broadband "noise," and bird song is often analyzed into themes and phrases, where researchers try to interpret the meaning or function of the different phrases. Marler and Tamura (1964) and Marler and Doupe (2000) believe that certain parts of the song contain certain types of information and that birds decode the songs. Emlen (1972) experimentally modified the songs of male indigo buntings (*Passerina cyanea*), and based on responses to playbacks, could identify the meaning of certain elements in the song (Fig. 11.12).



**Fig. 11.12** Male indigo bunting (*Passerina cyanea*) produces a song where certain elements of the song provide meaning to the listener. Photo "IndigoBuntingonPlant.jpg" by Kevin Bolton; https://wordpress.org/openverse/image/15bcd71f-0728-4bda-8122-38fcf4a82ce6/. Licensed under CC BY 2.0; https://creativecommons.org/licenses/by/2.0/

The male humpback whale is a well known marine singer. Males within each population of whales sing the same song, but each population of whales has its own unique song (rather like a dialect), which can sound different from the song in other populations. Within each population, the song structure changes gradually over the mating season and between years. A call unit can drop out of the repertoire, be replaced with another unit, or units can be added. These changes are known as song evolutions, as the song structure evolves gradually within a

population over time. Songs can also completely change between 1 year and the next, known as a song revolution. This is thought to be due to the influx of males from a different population, carrying with them their own song. Males from the original population then pick up and learn this new song causing the song within that population to completely change (Noad et al. 2000).

A duet is an exchange of sounds or substrate-borne vibrations between a pair of animals often produced in rapid succession (Fig. 11.13). The duet may be so rapid, that it is difficult to distinguish which animal is producing the various parts. It functions as a contact-maintaining signal and individual mated pairs within a species can develop their unique duet helping them to maintain contact with their partner. Duets are especially common in frogs, birds (cranes, sea eagles, geese, quail, grebes, woodpeckers, barbets, megapode scrub hens, kingfishers, ravens, cuckoo-shrikes, and honey-eaters), tree shrews (mammalian order Scandentia), and siamang (*Symphalangus syndactylus*), as well as being common in major groups of insects that communicate via substrate-borne vibrations. Species that perform duets often are monogamous (such as siamangs) and the two sexes resemble each other in appearance (that is, they are not dimorphic).

Duets are used when mated pairs are required to remain in touch over long periods of time. Duetting can be especially important within environments, such as in dense vegetation, where birds cannot see each other. By duetting, pairs keep close to each other, and in synchrony, so when conditions in a variable environment become right, mating can be achieved quickly and efficiently. In most gibbon species (family Hylobatidae), males, and some females, sing solos that function to attract mates and advertise their territory. If a male and female like one another's song, they will find each other and conduct a short mating dance followed by a long vigorous mating ritual. The song dialect is used to identify the singing gibbon's species and the area it is from. Therefore, duetting also reduces hybridization with closely related species (Mitani and Marler 1989).

### 11.6.7 From Chorusing to Copulation

Males that chorus (e.g., frogs, toads, and insects such as locusts (order Orthoptera) and cicadas (order Hemiptera)), attract females to a localized area. A classic example of this are the periodical cicadas (*Magicicada* sp.). Millions of 17-year cycle cicada gather to mate in forests in the eastern United States. Males aggregate into chorus centers and attract mates by producing high-intensity sounds (Fig. 11.13). The desert locust



**Fig. 11.13** A duet of ravens (*Corvus corax*). Photo "Ravens' Duet" by Ron Mead; https://www.flickr.com/photos/14093853@N04/2678807340 . Licensed under CC BY 2.0; https://creativecommons.org/licenses/by/2.0/



**Fig. 11.14** Desert locusts (Acrididae) emerge and go into flight *en masse*. Photo "Locust" by [nivs]; https://www.flickr.com/photos/42805979@N00/34263361. Licensed with CC BY-SA 2.0; https://creativecommons.org/licenses/by-sa/2.0/

(*Schistocerca gregaria*) forms one of the most intense swarms (Fig. 11.14), and can be found in countries such as Kenya, Somalia, India, and Saudi Arabia. Their loud chorusing is a means of sexual advertisement. BBC News reported on the "biblical locust plagues of 2020", when these insects swarmed in large numbers in East Africa (BBC News 2020).

The gecko *Ptenopus garrulus* produces loud continuous chirruping during a dusk chorus (Walker, 1998). These calls strengthen social bonding during sexual and courtship activities and are often produced together with visual and tactile behaviors.

An example of a more spatially contained event used by male sage grouse (*Centrocercus urophasianus*) to attract mates acoustically and visually is leks. Male sage grouse form large courtship leks in a social arena to produce elaborate visual displays with their gular pouches and the accompanying sounds of "swish-swish-coo-oo-poink" (Fig. 11.15; Bush et al. 2010). This study (p. 343) found that despite lekking behavior, male–male competition was spread out spatially and females often covered the entire social arena before copulating. Leks also are increasingly being recognized in invertebrates that communicate through substrate-borne vibrations, such as the prairie mole cricket (*Gryllotalpa major*). In this species, a male stridulates from inside a burrow he constructs in the soil, producing an airborne (sound) component that signals to fly females as a sexual advertisement. The same stridulation event has a substrate-borne component (vibration) that is used by nearby males to aid in spacing their burrows (Hill 1999).



**Fig. 11.15** Male Greater Sage-Grouse (*Centrocercus urophasianus*) by USFWS Pacific Southwest Region; https://www.flickr.com/photos/54430347@N04/6928668188. Licensed under CC 2.0; https://creativecommons.org/licenses/by/2.0/

After mate attraction, comes copulation. Ovulation in female alpacas (*Vicugna pacos*) is thought to be simulated during copulation, where the male produces a loud "orrgle" for 30 to 45 minutes while mounting the female (Abba et al. 2013). Even after copulation, calling may continue, where the tree frog *Phyllomedusa* (Hylidae) gives a separate call after oviposition.

## 11.7 Comparing Human Language to Nonhuman Auditory Communication

Despite the phenomenal array of different types of auditory communication in the different species, what are the defining characteristics of human language? Human language involves the use of vocal sounds that are symbolic of meanings, and therefore context independent. Thus, human language can be understood in the total absence of the communicator, such as when written, or when heard on the telephone.

There is a vast literature on human language, and a whole field of study: linguistics. Many scientists believe that the development of human language was the most important evolutionary step in distinguishing humans biologically. It is also widely maintained that development of human language was responsible for the further cognitive development of humans. Interestingly, nonhumans respond to general sounds and emotions in human language. More recent work has shown that some primates, dogs, marine mammals, horses, and elephants comprehend individual words and phrases. In fact, with experience, they understand a great deal more human language than we previously assumed (e.g., de Waal 2016; Kiley-Worthington 2017). Young human or nonhuman mammals do not only learn the meaning of words by conditioning as the behaviorists believed (Skinner 1957), but they also learn by observing others, imitation, and learning about cause and effect.

One of the first experiments to test if nonhumans could learn to speak a human language was the Kelloggs' studies (Kellogg and Kellogg 1933). This family raised a young chimp *Pan troglodytes* with their son and treated her similarly. At the end of several years, although their son was talking, the chimp found great difficulty making human sounds, and managed only "mama." The conclusion was that the chimp's inability to learn language implied that chimps have lower intelligence than humans. However, later it was discovered that the reason for her difficulty in making speech sounds was not a mental/cognitive lapse, it was physiological. She did not have the necessary muscles to control the sophisticated movements of the tongue, larynx, buccal and nasal cavities in order to make the different sounds (Lyn 2012). More recently, Fitch (2011) has argued that humans have what he called a "language ready brain." However, Savage-Rumbaugh et al. (2009) argue strongly that human language may not be any more sophisticated than ape languages. This is supported by the recognition of the many mental homologies between humans and other mammals (e.g., Kiley-Worthington 2017).

Since the middle of the twentieth century, the distinguishing features found in human language have been widely discussed, and the synopsis developed by Hockett (1960) is still widely adhered to. The first question is to what degree these defining features are found in other species (Table 11.2).

This list has been elaborated, extended, and modified, to include tactile, visual, taste, and olfactory communication (e.g., Christin 1999). The vocal repertoire of many species has been shown to fulfill most of these characteristics, and a list of some of the most pertinent studies is given here (e.g., Fitch 2011; Herman et al. 1984; Schusterman and Kastak 1998; Nehaniv and Dautenhahn 2002; Rendell and Whitehead 2001; Christiansen and Kirby 2003).

To simplify the differences between human spoken language, and communication attributes of other species, there are two human specializations. The first is that the human spoken language, unlike auditory communication of many other species (although not all), is mainly (but not exclusively) context independent. That is, the same word means the same thing in any context. Humans have developed this

**Table 11.2** Design features of human language and whether they have been recorded in other species. The species listed here are only examples, since there are others for which better evidence exists

| Design features | Humans | Chimpanzees | Horses | Elephants |
|---|---|---|---|---|
| PRODUCTIVITY<br>Different components together at different times | + | + | + | + |
| ARBITRARINESS<br>Different responses to same display | + | + | + | + |
| INTERCHANGEABILITY<br>One display triggers another | + | + | + | ? |
| SPECIALIZATION<br>Not directly related to consequences | + | + | + | + |
| DISPLACEMENT<br>Key features not related to antecedents | + | + | + | + |
| CULTURAL TRANSMISSION<br>Differences between populations as a result of learning | + | + | + | + |
| DUALITY<br>Symbols form sentences; components of expression contribute to whole interpretation | + | + | ? | ? |

characteristic much further than other species, and as a result, the meaning of what they are saying can be assessed whatever the situation, whether it be on the telephone, read, or written. However, it is true that many words can have multiple meanings or are used in specific contexts. Furthermore, using the same word in different communication contexts can change its meaning. Meanwhile, primate alarm calls seem to share a lot of features of words. The other important characteristic is that human language is highly symbolic. Again, this is not a unique characteristic of human language. For example, movements such as a horse swishing his tail, which may mean he will kick you, and ritualized displays, such as the courtship preening of Mandarin ducks (*Aix galericulata*; Fig. 11.16) are also highly symbolic. However, humans have taken symbolism further so that symbols can be built on top of each other. For example, one dog can be seen to be a dog and only one, but it can also be represented by a 1. Another 1 can be added, which is represented as 2. This led to the emergence of mathematics, and to further symbolic links in formulae culminating in our explanations of gravity or electricity and other phenomena in the world.

Some research has concentrated on teaching apes and marine mammals to develop and use a language that has features characteristic of human language. This includes teaching chimpanzees sign languages, and more recently, to use computer symbols. Interestingly Washoe, one of the first chimps, was taught American Sign Language. This chimp eventually managed to combine symbols to produce new meanings. For example, when asked what a duck was when swimming in the water, she signed it was a



**Fig. 11.16** Mandarin ducks (*Aix galericulata*) perform a specialized courtship routine. The males shake and bob their heads, as well as mocking drinking and preening, while raising their crest and orange sail feathers to "show off." They also incorporate sound into their courtship in the form of a whistling call. "Mandarin duck" by Tambako the Jaguar; https://www.flickr.com/photos/8070463@N03/853400195. Licensed under CC BY-ND 2.0; https://creativecommons.org/licenses/by-nd/2.0/

"water bird" (Gardner and Gardner 1984). Gluck (2016), in his account of grappling with central philosophical problems in animal ethics, recollects one of his weekly lab meetings (he was part of a research lab known for numerous breakthroughs in psychology and animal behavior) where the graduate students would discuss their research and topics of the day; signing chimps was a hot topic at the time. He noted that one of the students, a bit of a maverick, inquired whether the chimp ever asked "Can I go home now?" or "Can I leave?" Gluck and the other students dismissed this as foolhardy and would spend the next two decades exploring how primate models could inform human biomedical and behavioral science. But that is still the question of our time. If a captive animal could, would they ask to be released? Would they ask "Why are you doing this to me?" These animal-intensive tests came under extreme criticism from other scientists (Terrace 1985). Since then, a gorilla, bonobos (*Pan paniscus*), and other chimps, have learned to use computer symbols as a human-type language (Hopkins and Savage-Rumbaugh 1991). Kenneally explored the origin of the first word, and speculated on which great apes might have been capable of speaking the first word. Among other things, she said that such a speaker would have to have the anatomical and physiological capacity for speech, but they would also have to have something to say. In her view, this probably eliminated chimps, which she thought were immature and lacking in focus, rather than cognitively limited (Kenneally 2007).

Thomas Nagel's (1974) thought-provoking question "What is it like to be a bat?" argues that humans might imagine what it is like to be another being but can never know the conscious mental state to be that species, or even another human. We can look at systems, patterns, and responses, but each species and every human retain their own secrets and have their own experiences. That does not mean we should not try to understand nonhuman auditory and vibrational communication signals. These different world views, or knowledge of the world, lead us to a study of the epistemology of different species. Let us hope that we begin seriously to investigate this before it is too late and many species have become extinct due to our actions, most of which are the consequences of human language.

## 11.8 Summary

With modern technological aids and further studies, the study of acoustic and substrate-borne vibrational communication has advanced considerably since Busnel's (1963) seminal work. The origins of acoustic communication are likely to be from sounds associated with moving about in the environment and breathing in and out through respiratory passages. These sounds have become specialized for communication. Likewise, as animals move, regardless of how quietly, the motions lead to vibrations through the substrate that can be detected by others of the same or different species. Responses to these vibrations by others are reinforced or are lethal to the receiver, but likely also inform the sender. The first step is for the sounds or vibrations to become ritualized, leading to displays. The development of the necessary sending and receiving structures, such as the larynx or the insect tymbal, and a sensory apparatus such as the ear or subgenual organ, facilitated the evolution of an extremely diverse range of auditory and vibratory signals and cues, of which only some are described here.

Auditory and vibratory communication each has advantages and disadvantages. Though a signal can travel through substrates, meaning the signaler does not have to be in visual range, it can be overheard by others. Atmospheric conditions can influence the signal and other sounds/vibrations can mask it. Geographic separation of animals within a population can cause auditory and vibrational signals to evolve over time into different dialects and cultural waves. This variation can eventually separate animals within a species into different populations. One thing that is becoming increasingly clear is that there is not much time to uncover more about the complexities of auditory and substrate-borne vibrational communication in nonhumans before

the behavior of our species, as human language users, has led to the extinction of many species.

## References


Abba MA, Bianchi C, Cavilla V (2013) Chapter 15—South American Camelids. In: Tynes VV (ed) The behavior of exotic pets. Wiley-Blackwell, New York

Amos JN, Harrisson JA, Radford JQ, White M, Newell G, Nally NM, Sunnucks P, Pavlova A (2014) Species-and sex-specific connectivity effects of habitat fragmentation in a suite of woodland birds. Ecology 95(6): 1556–1568. https://doi.org/10.1890/13-1328.1

Aubin T, Jouventin P (1998) Cocktail-party effect in king penguin colonies. Proc R Soc B 265:1665–1673

Bain DE, Dahlheimm ME (1994) Effects of masking noise on detection thresholds of killer whales. In: Bain DE, Dahlheim ME (eds) Marine mammals and the Exxon Valdez. Elsevier, New York

Baptista LF, Gaunt SSL (1997) Social interaction and vocal development in birds. In: Snowden CT, Hausberger M (eds) Social influences on vocal development. Cambridge University Press, Cambridge

Baze W (1950) Just elephants. Corgi Books, London

Bee MA (2007) Sound source segregation in grey treefrogs: spatial release from masking by the sound of a chorus. Anim Behav 74(3):549–558. https://doi.org/10.1016/j.anbehav.2006.12.012

Bennet-Clark H (1998) How cicadas make their noise. Sci Am 278(5):58–61. http://www.jstor.org/stable/26057783. Retrieved 7 Feb 2021

Bennet-Clark HC (2000) Resonators in insect sound production: how insects produce loud pure-tone songs. J Exp Biol 202:3347–3357

Benney KS, Braaten RF (2000) Auditory scene analysis in estrildid finches (*Taeniopygia guttata* and *Lonchura striata domestica*): a species advantage for detection of conspecific song. J Comp Psychol 114:174–182

Bradbury JW, Vehrencamp SL (1998) Principles of animal communication. Sinauer Associates, Sunderland, MA

British Broadcasting Corporation (BBC) News (2020) The biblical locust plagues of 2020. https://www.bbc.com/future/article/20200806-the-biblical-east-african-locust-plagues-of-2020

Brownell P, Farley RD (1979) Orientation to vibrations in sand by the nocturnal scorpion *Paruroctonus mesaensis*: mechanisms of target localization. J Comp Physiol A 131:31–38

Brumm H, Todt D (2002) Noise-dependent song amplitude regulation in a territorial songbird. Anim Behav 63(5):891–897. https://doi.org/10.1006/anbe.2001.1968

Brumm HB, Zollinger SA (2011) The evolution of the Lombard effect: 100 years of psychoacoustic research. Behav 148(11–13):1173–1198. https://doi.org/10.1163/000579511X605759

Brumm H, Schmidt R, Schrader L (2009) Noise-dependent vocal plasticity in domestic fowl. Anim Behav 78:741–746

Bush K, Aldridge CL, Carpenter JE, Paszkowski CA, BoyceM CDW (2010) Birds of a feather do not always Lek together: genetic diversity and kinship structure of greater sage-grouse (*Centrocercus urophasianus)* in Alberta. Auk 127(2):343–353. https://doi.org/10.1525/auk.2009.09035

Busnel RG (ed) (1963) Acoustic behavior of animals. Elsevier, New York

Carter GG, Wilkinson GS (2016) Common vampire bat contact calls attract past food-sharing partners. Anim Behav 116:45–51. https://doi.org/10.1016/j.anbehav.2016.03.005

Catchpole CK, Slater PJB (2008) Bird song. Biological themes and variations, 2nd edn. Cambridge University Press, Cambridge

Charlton BD, Reby D, McCombe K (2007) Female perception of size-related formant shifts in red deer (*Cervus elaphus*). Anim Behav 74:707–714

Christiansen MH, Kirby S (2003) Language evolution. Oxford University Press, New York

Christin AM (1999) Les origines de l'écriture: Image, signe, trace. Le Débat 106(4):28. https://doi.org/10.3917/deba.106.0028

Clay Z, Smith CL, Blumstein DT (2012) Food-associated vocalizations in mammals and birds: what do these really mean? Anim Behav 83:323–330

Cocroft R, Gogala M, Hill PSM, Wessel A (2014) Studying vibrational communication. Springer, Berlin

Cooper LD, Randall JA (2007) Seasonal changes in home ranges of the giant kangaroo rat (*Dipodomys ingens*): a study of flexible social structure. J Mammol 88:1000–1008. https://doi.org/10.1644/06-MAMM-A-197R1.1

Cynx J, Lewis R, Tavel B, Tse H (1998) Amplitude regulation of vocalizations in noise by a songbird, *Taeniopygia guttata*. Anim Behav 56(1):107–113. https://doi.org/10.1006/anbe.1998.0746

David Mech LH, Boitani L (eds) (2003) Wolves: behavior, ecology, and conservation. University of Chicago Press, Chicago, p 472

De Waal F (2016) Are we smart enough to know how smart animals are? W W Norton & Company, New York

Dent ML, Larsen ON, Dooling RJ (1997) Free-field binaural unmasking in budgerigars (*Melopsittacus undulatus*). Behav Neurol 111(3):590–598. https://doi.org/10.1037/0735-7044.111.3.590

Dunlop RA (2017) Potential motivational information encoded within humpback whale non-song vocal sounds. J Acoust Soc Am 141(3):2204–2213. https://doi.org/10.1121/1.4978615

Dunlop RA, Cato DH, Noad MJ (2010) Your attention please: increasing ambient noise levels elicits a change in communication behaviour in humpback whales (*Megaptera novaeangliae*). Proc R Soc B 277(1693): 2521–2529. https://doi.org/10.1098/rspb.2009.2319

Dunlop RA, Cato DH, Noad MJ (2014) Evidence of a Lombard response in migrating humpback whales (*Megaptera novaeangliae*). J Acoust Soc Am 135(1): 430–437

Emlen ST (1972) An experimental analysis of the parameters of bird song eliciting species recognition. Behaviour 41(1/2):130–171

Endler JA (2014) The emerging field of tremology. In: Cocroft R, Gogala M, Hill PSM, Wessel A (eds) Studying vibrational communication. Springer, Berlin, pp vii–vix. https://doi.org/10.1007/978-3-662-43607-3

Evans CS, Marler P (1991) On the use of video images as social stimuli in birds – audience effects on alarm calling. Anim Behav 41:17–26

Evans CS, Marler P (1994) Food calling and audience effects in male chickens, *Gallus gallus* – their relationships to food availability, courtship and social facilitation. Anim Behav 47(5):1159–1170

Ey E, Pfefferle D, Fischer J (2007) Do age- and sex-related variations reliably reflect body size in non-human primate vocalizations? A review. Primates 48:253–267

Farabaugh SM, Dooling RJ (1996) Acoustic communication in parrots: laboratory and field studies of budgerigars, *Melopsittacus undulatus*. In: Kroodsma DE, Miller EH (eds) Ecology and evolution of acoustic communication in birds. Cornell University Press, New York, pp 97–117. https://doi.org/10.7591/9781501736957

Fischer J, Kitchen D, Seyfarth RM, Cheney DL (2004) Baboon loud calls advertise male quality: acoustic features and relation to rank, age, and exhaustion. Behav Ecol Sociobiol 56:140–148

Fitch WT (1997) Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. J Acoust Soc Am 102:1213. https://doi.org/10.1121/1.421048

Fitch WT (2011) Unity and diversity in human language. Philos Trans R Soc Lond B Biol Sci 366(1563): 376–388. https://doi.org/10.1098/rstb.2010.0223

Gannon WL, Lawlor TE (1989) Variation of the Chip vocalization of three species of Townsend chipmunks (Genus *Eutamias*). J Mammal 70(4):740–753. https://doi.org/10.2307/1381708

Gannon WL, O'Farrell MJ, Corben C, Bedrick EJ (2003) Call character lexicon and analysis of field recorded bat echolocation calls. In: Thomas J, Moss C, Vater M (eds) Echolocation in bats and dolphins. University of Chicago Press, Chicago, IL, pp 478–484

Gardner RA, Gardner BT (1984) A vocabulary test for chimpanzees (*Pan troglodytes)*. J Comp Psychol 98(4): 381–404. https://doi.org/10.1037/0735-7036.98.4.381

Garland EC, Goldizen AW, Rekdahl ML, Constantine R, Garrigue C, Daeschler Hauser N, Poole MM, Robbins J, Noad MJ (2011) Dynamic horizontal cultural transmission of humpback whale song at the ocean Basin Scale. Curr Biol 21(8):687–691. https://doi.org/10.1016/j.cub.2011.03.019

Gerhardt HC, Bee MA (2006) Recognition and localization of acoustic signals. In: Narins PM, Feng AS, Fay RR, Popper AN (eds) Hearing and sound communication in amphibians, vol 28. Springer, New York, pp 113–146

Gluck JP (2016) Voracious science and vulnerable animals. A primate scientist's ethical journey. University of Chicago Press, Chicago, IL

Gordon SD, Tiller B, Windmill JFC, Krugner R, Narins PM (2019) Transmission of the frequency components of the vibrational signal of the glassy-winged sharpshooter, *Homalodisca vitripennis*, in and between grapevines. J Comp Physiol 205:783–791

Henry L, Barbu S, Lemasson A, Hausberger M (2015) Dialects in animals: evidence, development and potential functions. Anim Behav Cogn 2(2):132–155. https://doi.org/10.12966/abc.05.03.2015

Herman LM, Richards DG, Wolz JP (1984) Comprehension of sentences by bottlenosed dolphins. Cognition 16(2):129–219. https://doi.org/10.1016/0010-0277(84)90003-9

Heyes C, Dickinson A (1990) The intentionality of animal action. Mind Lang 5(1):87–104

Hill PSM (1998) Environmental and social influences on calling effort in the prairie mole cricket (*Gryllotalpa major*). Behav Ecol 9(1):101–108. https://doi.org/10.1093/beheco/9.1.101

Hill PSM (1999) Lekking in *Gryllotalpa major*, the prairie mole cricket (Insecta: Gryllotalpidae). Ethology 105: 531–545

Hill PSM (2008) Vibrational communication in animals, 1st edn. Harvard University Press, London

Hill PSM, Wessel A (2016) Biotremology. Curr Biol 26: R181–R191

Hill PSM, Lakes-Harland R, Mazzoni V, Narins PM, Virant-Doberlet M, Wessel A (2019) Biotremology: studying vibrational behavior. Springer Nature, Cham, Switzerland

Hill PSM, Mazzoni V, Stritih Peljhan N, Virant-Doberlet M, Wessel A (2022) Biotremology: physiology, ecology and evolution. Springer Nature, Cham, Switzerland

Hoch H, Deckert J, Wessel A (2006) Vibrational signalling in a Gondwanan relict insect (Hemiptera: Coleorrhyncha: Peloridiidae). Biol Lett 2:222–224

Hockett CF (1960) The origin of speech. Sci Am 203:88–111

Hopkins WD, Savage-Rumbaugh ES (1991) Vocal communication as a function of differential rearing experiences in *Pan paniscus*: a preliminary report. Int J Primatol 12(6):559–583

Hulse SH, MacDougall-Shackleton SA, Wisniewski AB (1997) Auditory scene analysis by songbirds: stream segregation of birdsong by European starlings (*Sturnus vulgaris*). J Comp Psychol 111:3–13

Huxley JS (1966) A discussion of ritualization of behavior in animals and man. Philos Trans R Soc B 251:247–271

Jackson A, DeArment R (1963) The lesser prairie chicken in the Texas panhandle. J Wildl Manag 27(4):733–737. https://doi.org/10.2307/379848

Janik VM, Slater PJB (2000) The different roles of social learning in vocal communication. Anim Behav 60(1): 1–11

Keighley MV, Langmore NE, Zdenek CN, Heinsohn R (2017) Geographic variation in the vocalizations of Australian palm cockatoos (*Probosciger aterrimus)*. Bioacoustics 26(1):91–108. https://doi.org/10.1080/09524622.2016.1201778

Kellogg WN, Kellogg LA (1933) The ape and the child: a comparative study of the environmental influence upon early behavior. Hafner Publishing, New York and London

Kenneally C (2007) The first word: the search for the origins of language. Viking, New York

Kiley M (1969) The origin and evolution of some displays in canids, felids and ungulates with particular reference to causation. vol 1. Vocalisations, vol 2 Tail and ear movements. D. Phil University of Sussex

Kiley M (1972) The vocalisations of ungulates canids and felids with particular reference to their origin, causation and function. Zeit fur Tierpsychol 31:71–222

Kiley-Worthington M (2017) The mental homologies of mammals. Towards an understanding of another mammals world view. Animals 7(12):87.

Kondo W, Watanabe S (2009) Contact calls: information and social function. Jpn Psychol Res 51(3):197–208. https://doi.org/10.1111/j.1468-5884.2009.00399.x

Kroodsma DE (ed) (1982) Acoustic communication in birds. Production, perception and design features of sounds, vol 1. Academic Press, New York

Kroodsma DE, Miller EH (eds) (1982) Acoustic communication in birds. Song learning and its consequence, vol 2. Academic Press, New York

Lakes-Harlan R, Strauss J (2014) Functional morphology and evolutionary diversity of vibration receptors in insects. In: Cocroft RB, Gogala M, Hill PSM, Wessel A (eds) Studying vibrational communication. Springer, Berlin, pp 277–302

Lyn H (2012) Apes and the evolution of language: taking stock of 40 years of research. In: Vonk J, Todd K (eds) Shackelford Oxford handbook of comparative evolutionary psychology, Chapter 19. Oxford University Press, Oxford. https://doi.org/10.1093/oxfordhb/9780199738182.013.0019

Manabe K, Sadr EL, Dooling RJ (1998) Control of vocal intensity in budgerigars (*Melopsittacus undulatus*): Differential reinforcement of vocal intensity and the Lombard effect. J Acoust Soc Am 103:1190. https://doi.org/10.1121/1.421227

Marler P (1961) Logical analysis of animal communication. J Theor Biol 1:295–317. https://doi.org/10.1016/0022-5193(61)90032-7

Marler P, Doupe AJ (2000) Singing in the brain. Proc Natl Acad Sci 97(7):2965–2967. https://doi.org/10.1073/pnas.97.7.2965

Marler P, Tamura M (1964) Culturally transmitted patterns of vocal behavior in sparrow. Science 146:1483–1486

Mazzoni V, Eriksson A, Anfora G, Lucchi A, Virant-Doberlet M (2014) Active space and the role of amplitude in plant-borne vibrational communication. In: Cocroft RB, Gogala M, Hill PSM, Wessel A (eds) Studying vibrational communication. Springer, Berlin, pp 125–145

McKelvey EGZ, Gyles JP, Michie K, Barquín Pancorbo V, Sober L, Kruszewski LE, Chan A, Fabre CCG (2021) *Drosophila* females receive male substrate-borne signals through specific leg neurons during courtship. Curr Biol. https://doi.org/10.1016/j.cub.2021.06.002

Mitani JC, Marler P (1989) A phonological analysis of male gibbon singing behavior. Behaviour 109(1/2): 20–45

Morris D (1956) The function and cause of courtship ceremonies in L'instinct en le comportement des animeaux et des hommes. Fondation Singer-Polignac, Colloque International Sur L'instinct, Paris, pp 261–266

Morris D (1957) Typical intensity and its relationship to the problem of ritualization. Behaviour 11:12

Morton ES (1977) On the occurrence and significance of motivational-structural rules in some bird and mammal sounds. Am Nat 111:855–869

Morton ES (1982) Grading, discreteness, redundancy and motivational-structural rules. In: Kroodsma DE, Miller EH (eds) Evolution and ecology of acoustical communication in birds. Academic Press, New York, pp 183–212

Nagel T (1974) What is it like to be a bat? Philos Rev 83(4):435–450

Narins PM (1990) Seismic communication in anuran amphibians. Bioscience 40:268–274

Narins PM, Lewis ER (1984) The vertebrate ear as an exquisite seismic sensor. J Acoust Soc Am 76:1384–1387

Narins PM, Losin N, O'Connell-Rodwell CE (2009) Seismic and vibrational signals in animals. In: Squire LR (ed) Encyclopedia of neuroscience. Elsevier, Amsterdam, pp 555–559

Nehaniv C, Dautenhahn K (2002) Imitation in animals and artifacts. MIT Press, Cambridge, MA

Noad MJ, Cato DH, Bryden MM, Jenner MN, Jenner KCS (2000) Cultural revolution in whale songs. Nature 408: 537. https://doi.org/10.1038/35046199

O'Connell-Rodwell CE, Arnason BT, Hart LA (1997) Seismic transmission of elephant vocalizations and movement. J Acoust Soc Am 102:3124

O'Connell-Rodwell CE, Arnason BT, Hart LA (2000) Seismic properties of Asian elephant (*Elephas maximus*) vocalizations and locomotion. J Acoust Soc Am 108:3066–3072

O'Connell-Rodwell C, Guan X, Puria S (2019) Vibrational communication in elephants: a case for bone conduction. In: Hill PSM, Lakes-Harlan R, Mazzoni V, Narins PM, Virant-Doberlet M, Wessel A (eds) Biotremology: studying vibrational behavior. Springer Nature, Cham, Switzerland, pp 259–276

O'Farrell MJ, Corben C, Gannon WL (2000) Geographic variation in the echolocation call of the hoary bat (*Lasiurus cinereus*). Acta Chiropterol 2(2):185–196

Oliveira RF, McGregor PK, Latruffe C (1998) Know thine enemy: fighting fish gather information from observing conspecific interactions. Proc R Soc B 265(1401): 1045–1049

Ota N, Soma M (2022) Vibrational signals in multimodal courtship displays of birds. In: Hill PSM, Mazzoni V, Stritih-Peljhan N, Virant-Doberlet M, Wessel A (eds) Biotremology: physiology, ecology and evolution. Springer Nature, Cham

Peake TM, McGregor PK (2004) Information and aggression in fishes. Anim Learn Behav 32(1):114–121

Podos J, Warren PS (2007) The evolution of geographic variation in birdsong. In: Advances in the study of behavior. Elsevier, New York

Potash LM (1972) Noise induced changes in calls of the Japanese quail. Psychonomic Sci 26(5):252–254

Randall JA (1984) Territorial defense and advertisement by footdrumming in bannertail kangaroo rats (*Dipodomys spectabilis*) at high and low population densities. Behav Ecol Sociobiol 16:11–20. https://doi.org/10.1007/BF00293099

Randall JA, Lewis ER (1997) Seismic communication between the burrows of kangaroo rats, *Dipodomys spectabilis*. J Comp Physiol 181(5):525–531. https://doi.org/10.1007/s003590050136

Ratnam R, Feng A (1998) Detection of auditory signals by frog inferior collicular neurons in the presence of spatially separated noise. J Neurophysiol 80:2848–2859. https://doi.org/10.1152/jn.1998.80.6.2848

Rendell L, Whitehead H (2001) Culture in whales and dolphins. Behav Brain Sci 24(2):309–382. https://doi.org/10.1017/S0140525X0100396X

Reznikova Z (2019) Evolutionary and behavioural aspects of altruism in animal communities: is there room for intelligence? In: Evolution: cosmic, biological, and social. Almanac, Dublin

Riede T, Fitch WT (1999) Vocal tract length and acoustics of vocalization in the domestic dog, *Canis familiaris*. J Exp Biol 202:2859–2867

Riondato I, Gamba M, Tan CL, Niu K, Narins PM, Yang Y, Giacoma C (2021) Allometric escape and acoustic signal features facilitate high-frequency communication in an endemic Chinese primate. J Comp Physiol 207:327–336

Roberts L, Howard DR (2022) Substrate-borne vibrational noise in the Anthropocene: From land to sea. In: Hill PSM, Mazzoni V, Stritih-Peljhan N, Virant-Doberlet M, Wessel A (eds) Biotremology: physiology, ecology and evolution. Springer Nature, Cham

Sales G, Pye D (1974) Ultrasonic communication in animals. Chapman & Hall, London

Salinas-Melgoza A, Wright TF (2012) Evidence for vocal learning and limited dispersal as dual mechanisms for dialect maintenance in a parrot. PLoS One. https://doi.org/10.1371/journal.pone.0048667

Sanvito S, Galimberti F, Miller EH (2007) Vocal signaling of male southern elephant seals is honest but imprecise. Anim Behav 73:287–299

Savage-Rumbaugh S, Rumbaugh D, Fields W (2009) Empirical kanzi: The ape language controversy revisited. Skeptic 15(1):25–33

Schaller GB (1964) The year of the gorilla. University of Chicago Press, Chicago, p 304

Scheiber IBR, Weiß BM, Kingma SA, Komdeur J (2017) The importance of the altricial – precocial spectrum for social complexity in mammals and birds – a review. Front Zool 14:3. https://doi.org/10.1186/s12983-016-0185-6

Schusterman RJ, Kastak D (1998) Functional equivalence in a California Sea lion: relevance to animal social and communicative interactions. Anim Behav 55(5): 1087–1095. https://doi.org/10.1006/anbe.1997.0654

Sebeok T (1977) How animals communicate. Indiana University Press, Bloomington

Skinner BF (1957) Century psychology series. In: Verbal behavior. Appleton-Century-Crofts, New York. https://doi.org/10.1037/11256-000

Slater PJB (1981) Cultural evolution in chaffinch song: process inferred from micro and macro geographical. Biol J Linnaean Soc 42:135–147

Slater PJB (1986) The cultural transmission of bird song. Trends Ecol Evol 1(4):94–97

Slater PJB (1989) Bird song learning: causes and consequences. Ethol Ecol Evol 1(1):19–46. https://doi.org/10.1080/08927014.1989.9525529

Slobodchikoff CN, Perla BS, Verdolin JL (2009) Prairie dogs: communication and community in an animal society. Harvard University Press, Cambridge, MA

Smith WJ (1969) Messages of vertebrate communication. Science 165(3889):145–150

Strusaker TT (1966) Auditory communication among vervet moneys (*Cercopithecus aethiops*). In: Alternment SA (ed) Social communication among primates. Chicago University Press, Chicago, pp 281–384

Stumpner A, von Helversen D (2001) Evolution and function of auditory systems in insects. Naturwissenschaften 88:159–170

Šturm R, Rexhepi B, López Díez JJ, Blejec A, Polajnar J, Sueur J, Virant-Doberlet M (2021) Vibroscape – an overlooked world of vibrational communication. iScience, revision in review

Sullivan J, Demboski JR, Bell KC, Hird S, Sarver B, Reid N, Good JM (2014) Divergence with gene flow within the recent chipmunk radiation (Tamias). Heredity 11:185–194. https://doi.org/10.1038/hdy.2014.27

Sutton D, Nadler C (1974) Systematic revision of three Townsend Chipmunks (*Eutamias townsendii)*. Southwest Nat 19(2):199–211. https://doi.org/10.2307/3670280

Terrace HS (1985) Animal cognition: thinking without language. Philos Trans R Soc Lond B 308(1135): 113–128. https://doi.org/10.1098/rstb.1985.0014

Vannoni E, McElligott AG (2008) Low frequency groans indicate larger and more dominant fallow deer (*Dama dama*) males. PLoS One 3:e3113

Virant-Doberlet M, Čokl A (2004) Vibrational communication in insects. Neotrop Entomol 33:121–134
Virant-Doberlet M, Mazzoni V, de Groot M, Polajnar J, Lucchi A, Symondson WOC, Čokl A (2014) Vibrational communication networks: eavesdropping and biotic noise. In: Cocroft R, Gogala M, Hill PSM, Wessel A (eds) Studying vibrational communication. Springer, Berlin, pp 93–123
Volodin IA, Matrosova VA, Frey R et al (2018) Altai pika (*Ochotona alpina*) alarm calls: individual acoustic variation and the phenomenon of call-synchronous ear folding behavior. Sci Nat 105:40. https://doi.org/10.1007/s00114-018-1567-8
Walker SF (1998) Animal communication. In: Mey JL (ed) Concise encyclopedia of pragmatics. Elsevier, Amsterdam, pp 26–35
Wells KD, Schwartz JJ (2006) The behavioral ecology of anuran communication. In: Narins PM, Feng AS, Fay RR, Popper AN (eds) Hearing and sound communication in amphibians, vol 28. Springer, New York, pp 44–86
Wessel A, Mühlethaler R, Hartung V, Kuštor V, Gogala M (2014) The tymbal: evolution of a complex vibration-producing organ in the Tymbalia (Hemiptera excl. Sternorrhyncha). In: Cocroft R, Gogala M, Hill PSM, Wessel A (eds) Studying vibrational communication. Springer, Berlin, pp 395–444
Wiley RH (1983) The evolution of communication: Information and manipulation. In: Halliday TR, PJB S (eds) Animal behavior, volume 2, communication, vol 225. W. H. Freeman, New York, pp 156–189
Wilson-Henjum GE, Job JR, McKenna MF et al (2019) Alarm call modification by prairie dogs in the presence of juveniles. J Ethol 37:167–174. https://doi.org/10.1007/s10164-018-0582-8

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



# Echolocation in Bats, Odontocetes, Birds, and Insectivores

12

Signe M. M. Brinkløv, Lasse Jakobsen, and Lee A. Miller

## 12.1 Introduction

Echolocation, a term coined by Griffin (1944, 1958), is an active sensory system. Echolocating animals emit sound signals and perceive their surroundings by way of the returned echoes. Using this approach, echolocators can determine the direction and distance to an object, the type of object, and whether it is moving or stationary. Echolocation (also known as biosonar) is used by most bats, odontocetes (toothed whales), oilbirds, and some swiftlets to negotiate, respectively, night skies, deep waters, or dark caves. In addition, soft-furred tree mice use echolocation in darkness for orientation (He et al. 2021). These are all habitats characterized by limited visibility, likely a key evolutionary driver for echolocation. Echo feedback may also provide functional sensory abilities in shrews and tenrecs.

The discovery of echolocation traces back to Lazzaro Spallanzani's suggestion in 1794 that bats could "see" with their ears. Griffin (1944, 1958) verified this idea much later when he demonstrated that bats produce ultrasonic sounds to collect information about their surroundings and concluded that "echolocation is an eye-opening discovery about animal behavior."

Demonstrating echolocation behavior means showing that the animal uses echoes of their outgoing sounds to locate and identify objects in their path. Several robust protocols exist for assessing echolocation ability and capacity in terrestrial and marine animals (Griffin 1958; Norris et al. 1961). Echolocation and ultrasound are not inherently linked. Many animals echolocate by signals fully or partly composed of frequencies readily audible to humans, such as the clicks of some odontocetes, certain bat species, and birds. Conversely, many non-echolocating animals use ultrasonic sounds for intraspecific communication.

A primary advantage of echolocation is that it allows animals to operate and orient in uncertain lighting conditions. At the same time, information leakage is a primary disadvantage of echolocation. The signals used in echolocation are audible to many other animals, such as competing conspecifics, predators, and prey. The evolutionary arms race between echolocating bats and several families of insects sensitive to ultrasound is a classic example of predator–prey co-evolution (Miller 1983; Miller and Surlykke 2001). Some fishes (Alosinae) hear high-frequency sounds (Mann et al. 1997; Wilson et al. 2008), which could suggest similarly co-evolving sensory abilities between odontocetes and their fish prey (Wilson et al. 2013).

S. M. M. Brinkløv
Department of Ecoscience – Wildlife Ecology, University of Aarhus, Aarhus C, Denmark
e-mail: brinklov@ecos.au.dk

L. Jakobsen · L. A. Miller (✉)
Department of Biology, University of Southern Denmark, Odense, Denmark
e-mail: lasse@biology.sdu.dk; Lee@biology.sdu.dk

© The Author(s) 2022

C. Erbe, J. A. Thomas (eds.), *Exploring Animal Behavior Through Sound: Volume 1*,
https://doi.org/10.1007/978-3-030-97540-1_12

In this chapter, we review basic concepts about echolocation, the variety of animals known to echolocate, the main types of echolocation signals they use, and how they produce and receive those signals. The topic of perception by echolocating animals is beyond the scope of this chapter.

## 12.2 Characteristics of Echolocation Signals

Echolocating animals use two broad classes of sounds. Toothed whales, rousette bats, and birds generate broadband clicks produced at varying rates. The vast majority of bats, however, use tonal echolocation signals, characterized by longer duration and either a constant frequency or, more commonly, frequency modulation (FM; i.e., sweeping across several frequencies over time). With the exception of certain bat species, echolocating animals time their outgoing pulses so the echo from a previous pulse does not overlap with the next outgoing signal, especially during general orientation and searching for prey. This separation ensures that the strong outgoing signal does not mask the fainter returning echoes from the previous signal (Jen and Suga 1976; Kalko and Schnitzler 1989; Verfuss et al. 2009). Bats and odontocetes both show characteristic changes in echolocation behavior as they approach objects. Notably, most species in both groups adjust the sound emission rate to the distance of the target. The click rate increases as they approach objects and numerous species emit a terminal buzz (i.e., a series of pulses or clicks in rapid succession) during prey capture (Fig. 12.1). In bats, these temporal changes are accompanied by a change from narrow to wider bandwidths and lower to higher frequencies as they move from an open to a cluttered aerial environment or detect an airborne insect prey. Such pronounced, systematic changes have not been documented in oilbirds or swiftlets.

Echolocation signals are often much higher in amplitude than other sounds produced by animals. Amplitudes of bat echolocation signals are typically given at a reference distance of 0.1 m in front of the mouth or nostril. For whales and birds, source levels are referenced to a distance of 1 m in front of the animal. Source levels of bats are variable, but generally higher in aerial-feeding bats that fly and search for prey in the open sky (typically 100–130 dB re 20 μPa at 0.1 m). Bats that fly and forage in vegetation use lower-amplitude signals. Among these, the so-called "whispering bats" (e.g., slit-faced bats (Nycteridae), false vampire bats (Megadermatidae), and many New World leaf-nosed bats (Phyllostomidae)), emit echolocation sounds at about 65–70 dB re 20 μPa at 0.1 m (Jakobsen et al. 2013a). The source level of a dolphin's echolocation signal is several orders of magnitude greater than that of a bat's signal, primarily owing to the different properties of the two media (see next section) (Madsen and Surlykke 2014). Echolocation clicks of bottlenose dolphins (*Tursiops truncatus*) can reach source levels of 225 dB re 1 μPa at 1 m peak-to-peak (Au 1993, p. 78). Source levels of oilbirds (*Steatornis caripensis*) are around 100 dB re 20 μPa root-mean-square (rms) at 1 m (Brinkløv et al. 2017), corresponding to roughly 120 dB re 20 μPa at 0.1 m, which is comparable to estimates from many bat species. Little has been documented about the source levels of swiftlets, tenrecs, and shrews.

Bats and toothed whales both emit the acoustic signal energy in a focused beam, with specific vertical and horizontal transmission patterns, akin to an "acoustic flashlight" focused on a certain search area. The open mouth of a bat, or the nose in nasal-emitting bats, shapes the transmitted beam (Hartley and Suthers 1987, 1989), which is much broader than that of dolphins (Madsen and Surlykke 2014). The dolphin's melon transmits the outgoing echolocation signals with a slightly elevated vertical beam above the rostrum (Au 1993). There is no information on signal directionality from oilbirds or swiftlets.

## 12.3 Differences in Echolocation Signals in Air and Water

Only a few of the 71 known species of toothed whales are proven to use echolocation, but by inference probably all of them do (Culik 2011),

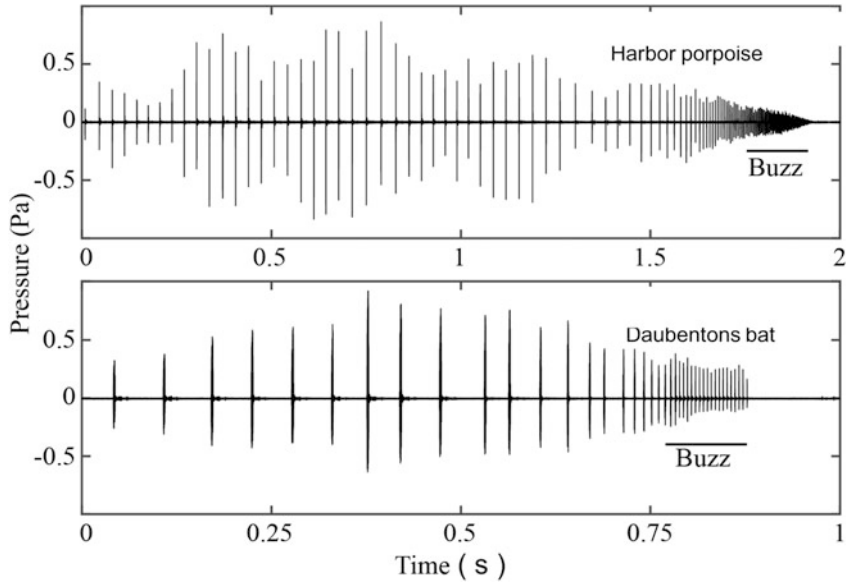



**Fig. 12.1** Echolocation sequence from a harbor porpoise (*Phocoena phocoena*) and a Daubenton's bat (*Myotis daubentonii*) as they approach and capture prey. Both species increase the rate of sound emission as they approach prey and emit a terminal buzz immediately before prey capture

as do presumably more than 1000 species of bats. For echolocators, there are three important differences between sound in air and sound in water: (1) density of the medium, (2) reflectivity of targets, and (3) maneuverability of the target (Madsen and Surlykke 2014). These differences severely influence the way echolocation has evolved in the two media (Au and Simmons 2007).

First, water is about 770 times denser than air: 1000 and 1.3 kg/m$^3$, respectively, partly explaining why sound travels about 4.4 times faster in water than in air (1520 m/s versus 344 m/s). For the same frequency of sound, the wavelength in water is about 4.4 times longer than in air. Longer wavelengths limit detection to larger targets because reflection depends on the relationship between the wavelength of the impinging sound and the size of the reflecting object (Urick 1983; also see Chap. 5, section on reflection). Sound at a given frequency reflects more effectively from smaller objects in air than in water. For example, the wavelength of a 100-kHz signal is 3.4 mm in air, and 15 mm in water. Thus, a sphere with a circumference greater than 3.4 mm strongly reflects the 100-kHz sound in air, while in water, the sphere must be larger than 15 mm in diameter.

The absorption coefficient (see Chaps. 5 and 6 on sound propagation) of the medium is a function of several factors, but frequency is the most important for echolocators. In seawater, the absorption coefficient for sound at 100 kHz is about 0.038 dB/m, while in air at the same frequency, it is much larger: 3.3 dB/m. In addition, sound pressure is lost through geometric spreading in both air and water. For spherical spreading, each time the distance is doubled, the sound pressure level of the emitted signal is halved (i.e., reduced by 6 dB). Taken together, sound absorption and geometric spreading mean that an echolocating dolphin can detect an object at much longer distances than can an echolocating bat (Madsen and Surlykke 2014).

Investigators often want to get a relative notion of the difference in amplitude of bat and dolphin echolocation signals. However, such a comparison should be done cautiously because of the different physical properties of air and water and the two different reference pressures. To compare

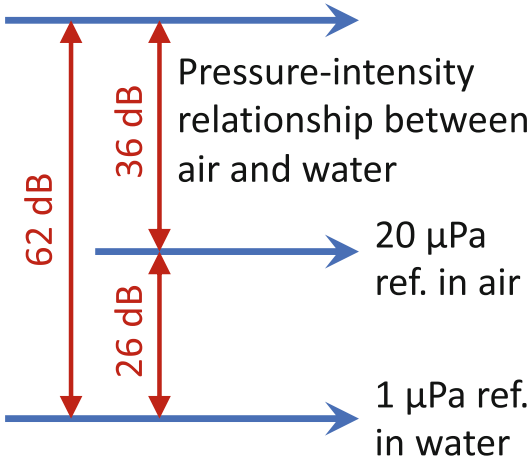



**Fig. 12.2** For sound sources of the same power or intensity, the sound pressure levels in air and water differ by 62 dB

a sound intensity level measured in dB in water to a reading in air, subtract 36 dB to compensate for the differences in acoustic impedance (i.e., density × sound speed; see Chap. 4, introduction to acoustics) between the two media. For the same source intensity, sound pressure in water is 60 times greater than in air (i.e., ~36 dB).

$$I_{\text{water}}/I_{\text{air}} = \left(p^2/\rho\, c\right)_{\text{water}}/\left(p^2/\rho\, c\right)_{\text{air}} = 1/3570$$

$$10\ \log_{10}(1/3570) = -36\ \text{dB}$$

where $p$ is sound pressure, $I$ is intensity, $\rho$ is density, $c$ is the speed of sound, and $\rho c$ is acoustic impedance. Then, subtract 26 dB ($20\ \log_{10}(20/1) = 26$ dB) to correct for the different reference pressures used for the decibel scales of sound in air and in water; i.e., 1 µPa in water and 20 µPa in air (Fig. 12.2). For example, if the sound pressure level of a dolphin click were 220 dB re 1 µPa (Au 1993), then a source with the same power would produce a click of 158 dB re 20 µPa in air ($220 - 36 - 26 = 158$ dB re 20 µPa), which is a very high sound pressure in air and well above the maximum sound pressure levels achieved by bats.

In air, there is a considerable difference in acoustic impedance between the medium and bat food, such as flying insects. There is, however, little impedance difference between seawater and toothed whale prey, such as fish or squid (Madsen et al. 2007). Accordingly, most sound from an echolocating toothed whale goes right through a fish or squid, producing low echo levels and making it difficult for the animal to detect its prey. In contrast, the air-filled swim bladders of some fish and hard features, such as the pen and beak of squid, reflect sound well, resulting in strong echoes.

In spite of substantial differences in the impedance and reflectivity of prey in air and in water, echo levels from airborne and aquatic prey are about the same. The target strength (TS) is the difference between the echo level (EL) measured 1 m from the target and the incident sound (IS) at the target: $TS = EL - IS$, where $EL$ and $IS$ are measured in dB re 20 µPa in air and 1 µPa in water, and $TS$ is in dB as the reference levels cancel out. Maximum target strength depends on the frequency of the echolocation signal and the reflectivity, size, and orientation of the prey with respect to incident sound. For cod, haddock, and saithe (400 to 500 mm long) the $TS$ (at 30 kHz) is −32 to −40 dB. For a moth (*Arctia caja*) with a 25–35 mm wingspan, $TS$ (at 20–50 kHz) is −42 dB; for the stonefly (*Plecoptera* sp.) with a wing-span of ~15 mm, $TS$ (at 10–37 kHz) is −47 dB (Miller 1983; Rydell et al. 1999). Despite more than a magnitude of difference in size, the target strengths of fish and insect prey are similar because of a combination of the differences in acoustic impedance of the medium and reflectivity of the prey.

Viscosity differences between air and water make toothed whales much less agile than bats. Toothed whales swim at about 2 m/s when capturing prey while bats fly at 2–10 m/s. After detection, a bat arrives at its prey much sooner than the toothed whale. A bat catching prey moves quickly because it is hardly hindered by friction from air. Bats typically take about a second to capture prey, while porpoises and dolphins need several seconds because the higher viscosity of water hinders their mobility. These differences occur despite similar ratios between body length

of predator and prey; a 3-m long dolphin is 6–15 times larger than its fish prey (20 to 50 cm long) and a 3–8 cm long bat is 5–10 times bigger than its insect prey. Bats often use their wing and tail membranes and even their feet to catch and manipulate insects. Toothed whales are streamlined with only pectoral and dorsal fins and flukes as appendages; they must catch and manipulate prey with their teeth and mouths (Miller 2010).

Despite very different selective pressures placed on bats and toothed whales, most of which are founded in the density and viscosity differences between air and water, they operate their biosonar in very similar ways. This similarity of the biosonar systems of bats and toothed whales (Fig. 12.5a) is a wonderful example of convergent evolution (Madsen and Surlykke 2014; Wilson et al. 2013).

## 12.4 Echolocation in Bats

Bats are the second-most species-rich order of mammals, currently comprising almost 1400 species (Burgin et al. 2018) and they play several trophic roles. Echolocating bats eat a diverse range of food including animals (insects, vertebrates), plant materials (leaves, fruit, nectar, and pollen), and even blood. The non-echolocating pteropodid bats all eat mainly plant materials. Traditionally, bats were arrayed in two suborders separating them into the echolocating Microchiroptera and the non-echolocating Megachiroptera, but recent phylogenetic studies do not support this division. Bats are now divided into Yinpterochiroptera and Yangochiroptera (Teeling 2009; Teeling et al. 2005). The non-echolocating pteropodid bats are found in the Yinpterochiroptera. This new division is intriguing because it creates two alternatives for the evolution of bat echolocation, either as a single event resulting in the loss of echolocation by the pteropodids or as two separate events. The current consensus favors a single origin of echolocation and subsequent loss in the pteropodids (Thiagavel et al. 2018; Wang et al. 2017).

### 12.4.1 Sound Production and Signal Characteristics

With the exception of the tongue-clicking *Rousettus* bats (10 species belonging to the pteropodid family), all ~1200 species of echolocating bats produce their echolocation signals in the larynx (Suthers and Hector 1988). The larynges and associated structures in bats are specialized to varying degrees from the basic mammalian pattern, notably the entire structure ossifies much earlier during development than in most mammals, and for many species the vocal tract and nasal passages are modified to filter frequencies used for echolocation (Au and Suthers 2014). Most echolocating bats emit sound through the open mouth, but bats in several families emit sound through the nostrils (Pedersen 1993). Bats emitting sound through the mouth generally have plain faces, while the bats emitting sound through the nose typically have elaborate structures surrounding the nostrils such as a nose-leaf that aids in sound radiation (Fig. 12.3).

The vast majority of echolocating bats are insectivorous. Most insectivorous bats hunt flying insects and typically vary the structure of their echolocation calls as they progress from searching to approaching and capturing prey. Traditionally, prey capture is divided into three phases (Fig. 12.4): a search, an approach, and a terminal phase (Griffin 1958; Griffin et al. 1960). In the search phase, bats emit long-duration, lower-frequency, narrowband signals (search calls) at a low repetition rate. After an object of interest is detected, the bats gradually reduce the duration and intensity of the signals; while they increase the rate and the bandwidth as they approach objects (approach calls). In the terminal phase, immediately before prey capture, the repetition rates may exceed 150 calls per second (the terminal buzz). Several reasons underlie these progressive changes in call emission. The search calls facilitate a long detection range as lower frequencies are attenuated much less than are higher frequencies (Lawrence and Simmons 1982b) and the long duration and narrow




**Fig. 12.3** Variation in bat facial morphology. (**a**) *Nyctalus noctula*, (**b**) *Murina cyclotis*, (**c**) *Plecotus auritus*, (**d**) *Mimon crenulatum*, (**e**) *Rhinolophus rouxii*, (**f**) *Hipposideros lankadiva*. Bats **a** and **b** are mouth emitting echolocators while **c**–**f** are nose emitters. Note that **c** does not have the associated nasal structures common in nose emitters. Photos by S. Brinkløv

bandwidth focus the energy of the call in a narrow range of the sensory system. These calls are, however, not ideal for accurate localization and object classification. Short-duration, broadband, high-frequency calls are much better suited for these tasks (Simmons et al. 1975). The switch from long-duration, narrowband, low-frequency calls in the search phase to short-duration, broadband, higher-frequency calls in the approach phase is a clear indication of object detection and it has been used to estimate detection distance in echolocating bats. However, it is important to note that this is a minimum measure as the bat may well have detected the object before adjusting its call parameters (Kalko and Schnitzler 1989, 1993).

Most echolocating bats, like toothed whales, emit an echolocation call and wait for echoes from objects of interest before emitting the next call (Madsen and Surlykke 2014). While this avoids perceptual errors associated with potentially assigning echoes to the wrong calls, it also means that the distance between the bat and objects of interest limits the call emission rate. As the bats approach an object, echoes return with progressively shorter delays and the bat can emit the calls at a higher rate, up to over 200 calls/s during the terminal buzz (Simmons et al. 1979, Fig. 12.4). While this is an impressively high call rate, the echoes are still received well before the next call is emitted. At the short distances between the bat and the prey when the buzz is emitted, the bat could theoretically increase the call rate to 1000 calls/s and still avoid call-echo ambiguity. Instead, the call rate is limited by the maximum speed of the superfast muscles that control each call emission (Elemans et al. 2011). Concurrent with the increase in call rate, the call duration decreases as distance to the object decreases. This is likely to prevent overlap

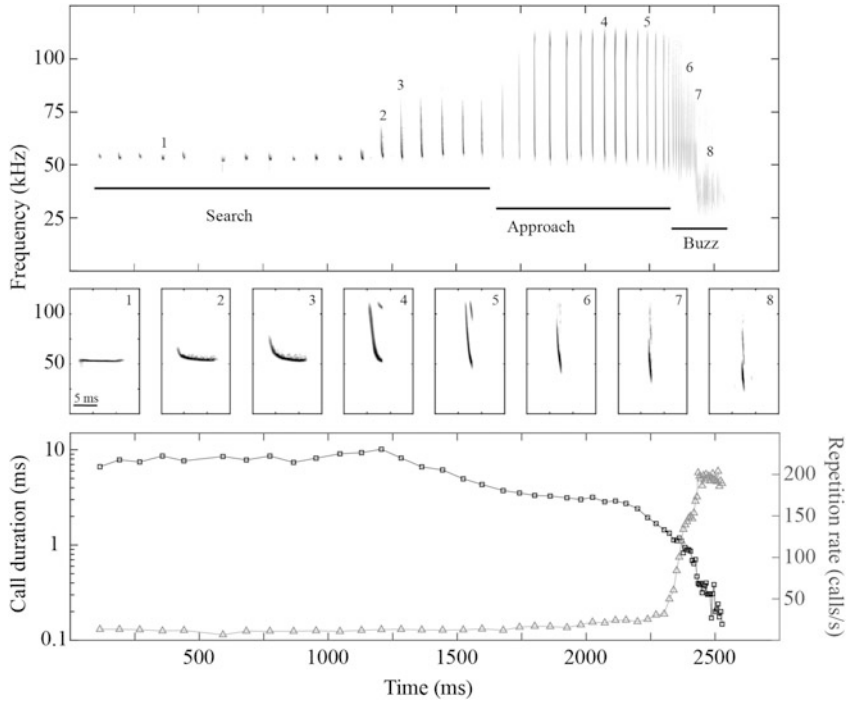



**Fig. 12.4** Echolocation call sequence emitted by a foraging soprano pipistrelle (*Pipistrellus pygmaeus*), illustrating the progressive change in call characteristics and emission rate as the bat searches for, approaches, and captures insect prey

between the emitted call and the returning echo since the much louder call emission will mask the quieter returning echo if the two overlap (Kalko and Schnitzler 1989, 1993). Hence, echoes from objects of interest are received in a clearly defined window between the end of call emission and the beginning of the next call. For example, a bat emitting calls of 8 ms duration at a call rate of 10 calls/s can resolve echoes from objects between 1.4 and 17 m distance without masking the returning echo during call emission and without the risk of call-echo ambiguity (Fig. 12.5).

While call rate and call duration define an overlap-free window, it is the energy and frequency of the emitted call together with the bat's hearing threshold and the nature of the echo-generating object that determine the range of the echolocation system. Echoes have to return with enough energy to be detected by the bat. Emitting more energy, either by increasing the intensity or duration of the call, increases the detection distance. Emitting lower frequencies also increases the detection distance because acoustic attenuation is less for lower frequencies. On the reflection side, small objects return quieter echoes and will therefore always be detectable at shorter ranges than large objects (Fig. 12.6). The structure and texture of the object also affects the level of the returning echo. Hard objects reflect more sound than soft objects and the same is true for plane or convex surfaces compared to concave surfaces (Urick 1983; also see Chap. 5, section on reflection). Additionally, the relationship between the wavelength of the sound impinging on the object and the size of the object affects how efficient the sound is reflected. If the wavelength becomes too long (i.e., the frequency too low) relative to the size of the object, very little sound is reflected (Fig. 12.6). This means that prey size imposes a lower frequency limit on bat echolocation (Houston et al. 2004; Pye 1993).

Bats are limited both physically and physiologically in how high a sound pressure they can produce. Supposedly, the main reason why they emit long-duration calls in the search phase is to increase the energy of the call. Emitting sound

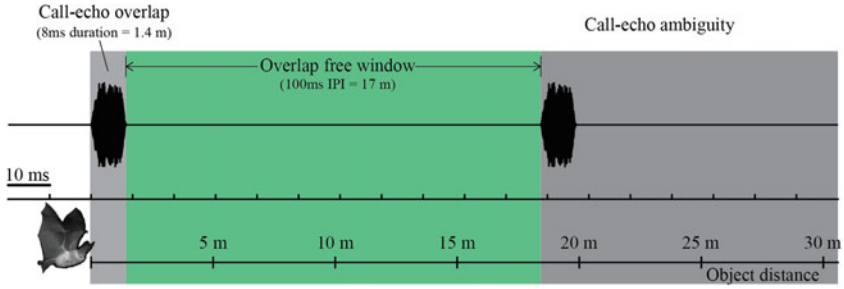



**Fig. 12.5** Schematic illustration of why most echolocating bats adjust call duration and call emission rate relative to target distance. Echoes received during call emission are masked by the louder call and echoes received after emission of the next call may create ranging ambiguity if assigned to the incorrect call. IPI: inter-pulse interval

directionally also increases the source level, that is the sound level measured directly in front of the animal. All bats studied to-date emit directional echolocation calls. Most bats increase their source level by 10 dB or more purely by focusing the sound as opposed to radiating sound equally in all directions (Jakobsen et al. 2013a). The highest source levels measured from bats are around 140 dB re 20 μPa rms at 0.1 m for the greater bulldog bat (*Noctilio leporinus*), but most reports of open-space aerial hawking bats are around 130 dB re 20 μPa rms at 0.1 m (Holderied et al. 2005; Hulgard et al. 2016; Surlykke and Kalko 2008). Combining knowledge of source level, signal frequency, hearing threshold, and the echo-generating object, the detection distance is relatively easy to estimate using a variation of the sonar equation (Urick 1983) (also see Chap. 6, section on the sonar equation):

$$RL = SL - 2 \times PL + TS$$

$$PL = 20 \times \log_{10}(\text{distance}/0.1\ \text{m}) + \alpha \times (\text{distance} - 0.1\ \text{m})$$

Here, *RL* is the received level, *SL* is the source level emitted by the bat, *PL* is the propagation (formerly, transmission) loss, $\alpha$ is the frequency-dependent attenuation in air, and *TS* is the target strength, a measure of how much sound is reflected from the object at 0.1 m relative to the sound impinging on the target. For an object to be detected by the bat, *RL* simply has to be above the bat's hearing threshold. The maximum distance that satisfies this requirement is the maximum detection distance. Estimated detection distances vary greatly between species, but it is clear that bat echolocation is a short-range system; the furthest estimates for large insect prey are around 10 m with most estimates below 5 m (Kalko and

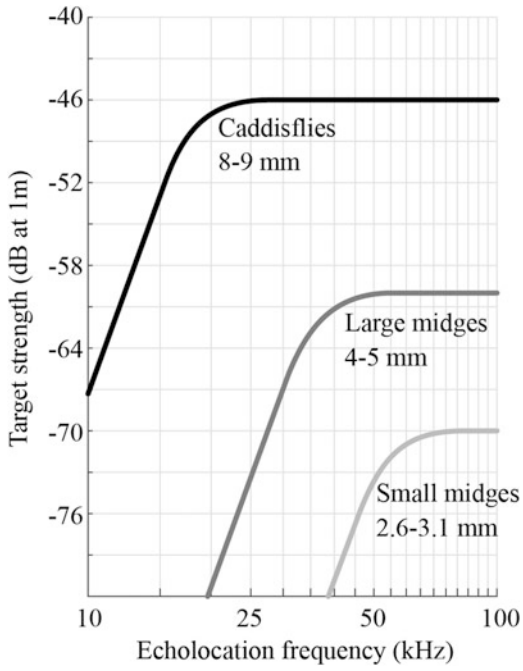



**Fig. 12.6** Target strength of three types of insect as a function of echolocation frequency illustrating how reflection depends on the relationship between object size and frequency. Smaller insects have lower target strength and require higher frequencies for efficient reflection. Indicated sizes are wing length. Based on data from Houston et al. (2004)

Schnitzler 1989, 1993; Nørum et al. 2012; Surlykke and Kalko 2008; Stilz and Schnitzler 2012).

The directional echolocation calls of bats allow an increased detection distance ahead of the bat while reducing the sound levels off to the sides and the back. This reduction in off-axis sound level offers an additional benefit as it reduces echoes from objects in these directions that are likely of little interest to the bats. Echoes from irrelevant objects are known as clutter echoes and reducing them simplifies the acoustic scene that the bats experience. The obvious disadvantage in emitting directional echolocation calls is the loss of echoes from relevant off-axis objects. The degree to which the benefits outweigh the costs of emitting a very directional echolocation call varies with the environment and the behavioral context. The directionality of the echolocation call is determined by the emitted frequency and the shape and size of the sound emitter. For mouth-emitting bats, this is the shape and size of the open mouth, and for nose-emitting bats, the shape and size of the nostrils and the nose-leaf (Hartley and Suthers 1987, 1989; Strother and Mogus 1970). Higher frequencies and larger emitters produce higher directionality (Fig. 12.7). Varying the frequency, shape, and size of the emitter allows the bats to adjust the directionality of the emitted call to suit their environment (Kounitsky et al. 2015; Surlykke et al. 2009b). During the final buzz of prey pursuit, bats can broaden their echolocation beam to increase peripheral echo levels and better track the prey (Jakobsen et al. 2015; Jakobsen and Surlykke 2010; Matsuta et al. 2013; Motoi et al. 2017). This is achieved in several species by a sudden drop in call frequency by nearly an octave (as illustrated in Figs. 12.4, 12.7, and 12.8) and is often referred to as the buzz II phase.

The majority of echolocating bats, and the focus of our description so far, hunt flying insects (aerial hawking bats) using relatively short-duration echolocation calls (also known as low duty-cycle calls, with duty cycle being the duration of the call divided by the time period (from the start of one call to the start of the next call). There are, however, many species that forage and echolocate differently. About 150 species, including the Old World horseshoe bats and hipposiderid bats (i.e., *Pteronotus parnellii* and closely related species in the family Mormoopidae from the New World), also feed on flying insects. These bats are so-called high duty-cycle echolocators and are able to broadcast and receive sound at the same time. While low

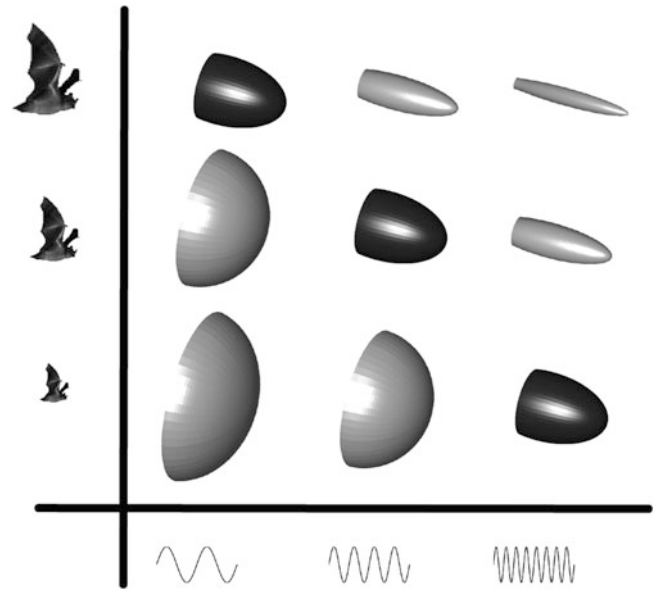


**Fig. 12.7** Echolocation call directionality as a function of emitter size and frequency. Directionality increases with increasing frequency and increasing size. Reprinted by permission from Springer Nature. Jakobsen L, Ratcliffe JM, Surlykke A. Convergent acoustic field of view in echolocating bats. Nature 493 (7430):93–96. https://www.nature.com/articles/nature11664. © Springer Nature, 2013b. All rights reserved

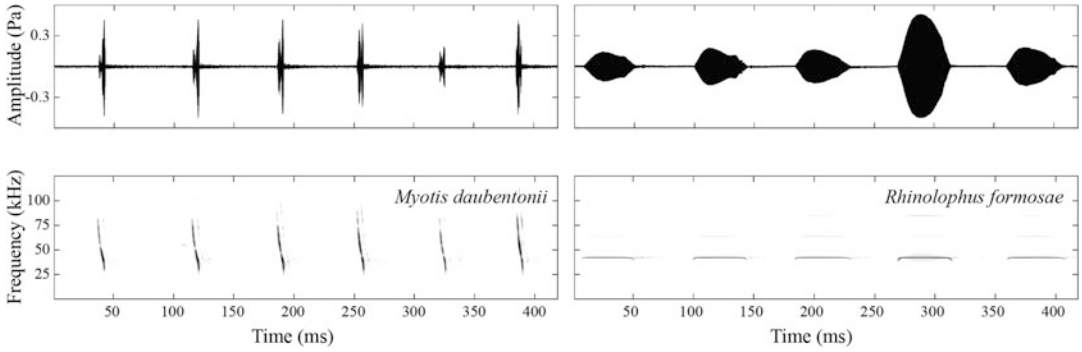



**Fig. 12.8** Echolocation calls emitted by a low duty-cycle bat (*Myotis daubentonii*) with strongly frequency-modulated calls (left) and a high duty-cycle bat (*Rhinolophus formosae*) with mostly constant frequency calls (right)

duty-cycle bats maintain a clear time separation between the emitted call and returning echo, high duty-cycle bats separate call and echo by frequency. They all emit much longer duration, constant-frequency echolocation calls with short intervals to navigate and forage (Fig. 12.8, Fenton et al. 2012). When an echo-generating object, such as a moth, moves relative to the bat, the echo returns to the bat at a slightly different frequency than the emitted call because of the Doppler shift. The classical example used to explain the Doppler shift phenomenon is the moving ambulance. When an ambulance moves toward a nearby listener, the siren appears to be higher in frequency than the one heard by someone riding in the ambulance, which does not change. The effect of Doppler shift is apparent when the ambulance passes and moves away from the listener. Now, the frequency abruptly changes from higher to lower in pitch. Doppler shift occurs because the speed of the moving ambulance is added to, or subtracted from, the speed of sound, raising or lowering the perceived pitch of the siren. The amount of the Doppler shift is doubled for echolocating animals, as the frequencies of both outgoing and returning signals are shifted. The Doppler shift experienced by an echolocating animal may be computed as:

$$\Delta f = (v_1 + v_2) \times f \times \cos\theta \times \frac{2}{c}$$

Here, $\Delta f$ is the amount of Doppler shift in Hz, $v_1$ is the speed of the echolocating animal in m/s, $v_2$ is the speed of the target in m/s (+ indicates movement away from the echolocator; − would be movement toward the echolocator), $f$ is the emitted frequency in Hz, θ is the angle in degrees between the echolocater and the target, and $c$ is the speed of sound in the medium (about 344 m/s in air and 1500 m/s in water).

Perception of a Doppler shift by an echolocator is facilitated by emitting long signals tuned to one frequency (narrowband or constant frequency) and by having acute hearing in the frequency band of the Doppler-shifted echo. Specifically, Doppler-shifted echoes are dominated by different frequencies than those dominating outgoing pulses (Fenton et al. 2012) and bats using this strategy are therefore not sensitive to overlap of the two.

Greater horseshoe bats (*Rhinolophus ferrumequinum*) detect the frequency and amplitude modulations of the Doppler-shifted echo from an insect to within a few Hz of the ~82 kHz carrier-frequencies of their echolocation calls (Neuweiler 2000). The bats that use Doppler-shifted echoes readily detect the wing beats of a fluttering insect and distinguish the prey from the background. Flutter-detection is a recurring theme among bats that exploit Doppler shifts (Goldman and Henson 1977; Schnitzler and Flieger 1983; Lazure and Fenton 2011).

Bats that exploit Doppler-shifted echoes are Doppler-shift compensators (DSC; Hiryu et al. 2016) because they continuously adjust the outgoing signal to ensure that the Doppler-shifted

echoes remain at the frequencies to which their acoustic foveae are tuned (Schuller and Pollack 1979, Schnitzler 1968; Schnitzler and Flieger 1983; Hiryu et al. 2016).

There is no current evidence that toothed whales or other echolocators using broadband clicks are capable of Doppler-shift compensation. However, the small harbor porpoise would be a good species to test for Doppler-shift sensitivity, as they have narrow auditory filters (Popov et al. 2006) and use relatively long clicks (100 μs) and narrowband echolocation signals centered around 130 kHz.

High duty-cycle bats, in general, have a highly specialized hearing to facilitate this type of echolocation and they modify their emitted echolocation calls such that the frequency of the returning echoes always falls within a very narrow frequency range for which their hearing is optimized (Fig. 12.8 and Sect. 12.4.2) (Schnitzler 1973; Schuller 1977). In spite of the large differences between high and low duty-cycle bats, the overall call emission pattern when catching flying insects is still remarkably similar. High duty-cycle bats still emit calls that correspond to the three phases of search, approach, and buzz when they pursue flying insects, including similar call-structure changes to those in the low duty-cycle bats: gradual source-level reduction, duration shortening, increasing repetition rate (Ratcliffe et al. 2013), and broadening of the echolocation beam during the terminal buzz (Matsuta et al. 2013).

Bats that do not forage for flying insects generally search for more conspicuous food. Many species hunt non-flying insects in dense vegetation, a strategy known as gleaning. Gleaning bats, in general, emit very short low-intensity calls that sweep over a broad range of frequencies (Denzinger and Schnitzler 2013). As noted earlier, such calls provide excellent localization and classification and the low intensities greatly weaken clutter echoes, which is particularly important when flying in dense vegetation. Fruit and nectar eating can be considered variations on the gleaning strategy, and the echolocation behavior of fruit-eating and nectar-drinking bats very closely resembles that of insect-gleaning bats (Denzinger and Schnitzler 2013). Notably, while these species often cluster their calls in groups with increased repetition rates when faced with increasing acoustic complexity, they do not emit the terminal buzz characteristic of bats that target flying insect prey (Gonzalez-Terrazas et al. 2016). In addition, they often rely on additional sensory input, such as olfactory cues (Gonzalez-Terrazas et al. 2016), or, in the special case of vampire bats, thermoreception (Kürten and Schmidt 1982).

### 12.4.2 Hearing Anatomy and Echolocation Abilities

The hearing of echolocating bats is based on standard mammalian hearing anatomy, including recognizable pinnae, tragus, ear canal, tympanic membrane, three middle ear bones, and a coiled cochlea. With few exceptions, they even have the same hearing threshold as most other mammals, measured at their best frequencies: 0 dB re 20 μPa (Fay 1988), Fig. 12.9. There are, however, notable specializations that relate to echolocation where bats differ from most mammals. It is clear that most bats have a larger than average pinna and tragus, but there is considerable variation across species in size and shape that likely relates to the bat's echolocation signals and foraging ecology (Coles et al. 1989; Obrist et al. 1993) (Fig. 12.3). In general, bats that complement their echolocation by passive listening for prey-generated sounds have larger pinnae than bats that rely solely on echolocation (Obrist et al. 1993). The pinna provides substantial directionality and acoustic gain depending on the relationship between pinna size and sound frequency. The pinnae of gleaning bats commonly amplify sound well below the bats' echolocation frequencies (Coles et al. 1989; Guppy and Coles 1988; Obrist et al. 1993; Schmidt et al. 1983). The acoustic gain provided by the large pinnae affords some bats extremely low hearing thresholds such as the impressive −20 dB re 20 μPa hearing threshold found in the brown long-eared bat (*Plecotus auritus*) and the Indian false vampire bat (*Megaderma lyra*) (Coles et al. 1989; Schmidt et al. 1983). While pinna structure plays a crucial

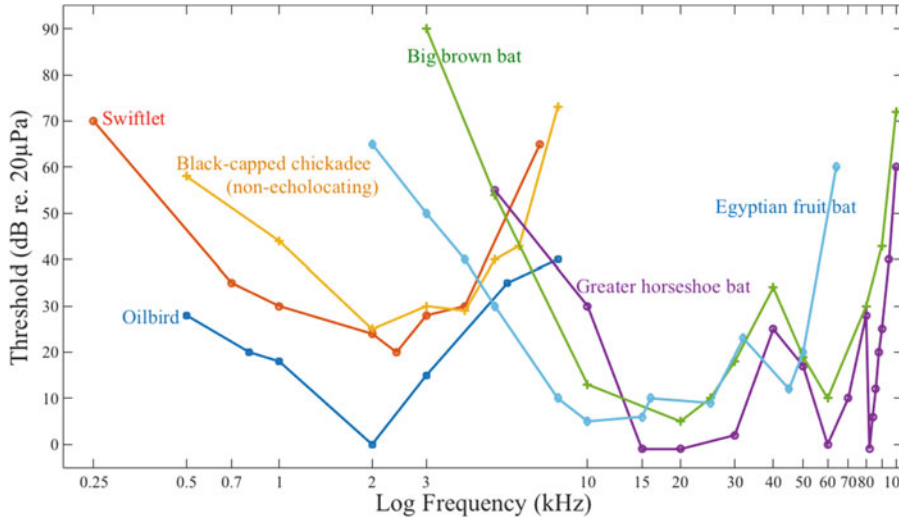



**Fig. 12.9** Audiograms of three echolocating bats and two echolocating bird species. A non-echolocating bird is shown for comparison. Bat thresholds are based on behavioral experiments, bird thresholds are derived from neurophysiological experiments. Green: big brown bat (*Eptesicus fuscus*, from Dalland 1965); light blue: Egyptian fruit bat (*Rousettus aegyptiacus*, from Koay et al. 1998); purple: greater horseshoe bat (*Rhinolophus ferrumequinum*, from Long and Schnitzler 1975); dark blue: oilbird (*Steatornis caripensis*, from Konishi and Knudsen 1979); red: swiftlet (*Aerodramus spodiopygia*, from Coles et al. 1987); yellow: black-capped chickadee (non-echolocating, from Wong and Gall 2015). Thresholds are not directly comparable between species due to differences in experimental conditions

role in bat echolocation, large external ears have a disadvantage during flight. Large ears create substantial drag, and it is likely that the ears of fast-flying bats are shaped as much by the aerodynamics of flight as by echolocation (Gardiner et al. 2008; Johansson et al. 2016; Vanderelst et al. 2015).

As mentioned above, bats decrease their emitted intensity progressively as they approach objects. This is primarily believed to function as gain control for the auditory system, a phenomenon also seen in echolocating odontocetes (see Sect. 12.5.2). If the bats kept their output level constant, the echo level would increase progressively by many orders of magnitude as the bat approached an object. Considering small insects as point sources, this increase would be $40 \times \log_{10}(r)$ or 12 dB per halving of distance $r$. So, the output call level generally decreases by 6 dB per distance halved (Boonman and Jones 2002; Brinkløv et al. 2013; Hartley 1992a, b; Lewanzik and Goerlitz 2018). Such a reduction results in a constant intensity at the object/prey, but a progressive increase in echo strength at the bat by +6 dB per halving of distance. However, the bat's auditory system reduces its sensitivity by an additional 6 dB per halving of distance, because as the bat vocalizes, the middle ear muscles contract to avoid self-deafening, increasing the bat's hearing threshold. This time-dependent change in hearing threshold corresponds almost perfectly to the missing 6 dB per halving of distance and presumably provides a constant perceived echo level for the bat (Hartley 1992a, b; Henson 1965; Suga and Jen 1975). The gradual relaxation of the middle ear muscles progressively decreases the bat's hearing threshold back to resting level. It is worth noting that this is under very predictable laboratory conditions and that in a real-life field scenario, the bats encounter much more unpredictable conditions and prey behavior. Recordings of prey capture in the field reveal that intensity reduction is much more variable and commonly exceeds 6 dB per halving of distance (Nørum et al. 2012). This subject is also

discussed below for harbor porpoises and dolphins.

Bat hearing is certainly specialized for echolocation and for high frequencies (Fig. 12.9). Other small mammals such as mice and rats have a similar high-frequency hearing. Bats are, however, much more sensitive up to their high-frequency limit and have very high sensitivity over a much wider range of frequencies. Comparing echolocating to non-echolocating bats, the cochlea is significantly larger relative to skull size, and the basilar membrane, where frequency coding occurs, is longer for echolocating bats compared to all other mammals (Kössl and Vater 1995). High duty-cycle bats have the longest basilar membranes containing an acoustic fovea, which is a large region of the membrane dedicated to a very narrow frequency range. The acoustic fovea provides the crucial frequency resolution and sharp tuning that allows high duty-cycle bats to separate call and echo by frequency instead of time (Bruns and Schmieszek 1980).

Bats use the time delay between their outgoing call and the returning echo to determine the distance to a target. They determine the horizontal direction to the object by comparing the input on the two ears. For bats, interaural intensity differences likely provide the main cues (Pollak 1988). The vertical direction is mainly coded by frequency-dependent reflections from the pinna and tragus (Lawrence and Simmons 1982a). Bats have excellent spatial resolution and accuracy. They consistently aim their echolocation beam to within less than 5° of their target both horizontally and vertically (Ghose and Moss 2003; Jakobsen and Surlykke 2010; Masters et al. 1985; Surlykke et al. 2009a) and can discriminate between two objects in the horizontal plane if they are more than 1.5° apart (Simmons et al. 1983) and, in the vertical plane, if they are more than 3° apart (Lawrence and Simmons 1982a).

Aerial hawking bats can easily be tricked into catching small pebbles thrown in the air. This is not because bats cannot distinguish pebbles from insects, but likely because most airborne items of a given size are edible to bats. Classification of small objects is based on temporal and spectral features of the echo generated by one or more reflections from the objects (Schmidt 1988; Simmons et al. 1990; Weissenbacher and Wiegrebe 2003), while the classification of large objects such as trees is more complex (Grunwald et al. 2004). The bat's resolution of a target depends on both the frequency of the emitted call (higher frequencies reflect more efficiently off smaller structures than do lower frequencies (Fig. 12.6 and Urick 1983) and the bat's ability to perceive these reflections. Bats are capable of distinguishing similar-sized objects with very minute textural differences. They can clearly distinguish small disks from mealworms when both are thrown in the air and smooth hanging beads from textured beads with the same overall echo-strength (Falk et al. 2011; Griffin et al. 1965).

Our account of bat echolocation only contains broad strokes. With around 1200 species of echolocating bats, the variation in echolocation design is vast, and while most follow the outline given here, there are many deviations and many bat species that utilize their echolocation in puzzling ways that are as yet unexplained.

## 12.5 Echolocation in Odontocetes

Among cetaceans, only species in the suborder Odontoceti (toothed whales) are known to echolocate (Au 1993). Bioacoustical research has focused on bottlenose dolphins, belugas, false killer whales, and killer whales (all in the families Monodontidae and Delphinidae) as well as porpoises (Phocoenidae), sperm whales (Physeteridae), and a few species of beaked whales (Ziphiidae).

Odontocetes use echolocation to orient in the aquatic environment, to detect, chase, and capture prey, and to socialize (Thomas et al. 2004; Thomas and Turl 1990). They have broadband hearing and a good ability to discriminate a signal in noise. Their echolocation signals have narrow beam patterns that can be modified, as can the amplitude and frequency content of outgoing clicks.

The bottlenose dolphin has been the "laboratory rat" of odontocete biosonar studies. A series

of experiments by US Navy researchers examined the ability of captive bottlenose dolphins (*Tursiops truncatus*) to detect subtle differences in human-made objects for military reconnaissance purposes (Au 1993, 2015; Moore and Popper 2019). They showed that dolphins wearing eyecups (so they could not see their targets) and using only echolocation could: (1) distinguish objects of the same shape, but of different materials (e.g., cylinders of glass, metal, or rock), (2) distinguish objects of the same material but different shapes (e.g., PVC cylinders, plates, squares, and tubes), (3) detect a 3-inch hollow metal sphere at about 115 m distance and a sphere of a few millimeters at a distance of about 50 m, (4) feed normally if blind, but if hearing-impaired become disoriented, (5) discriminate metal cylinder targets with different wall-thickness (difference as little as 0.00 1 mm), and (6) control the amplitude and frequency of their outgoing pulses, such that in areas of high ambient noise, they produced louder and higher-frequency pulses.

### 12.5.1 Sound Production and Signal Characteristics

Most dolphins emit whistles and burst-pulse sounds for intraspecific communication and brief broadband clicks for echolocation. Figure 12.10 shows four echolocation clicks from a false killer whale (*Pseudorca crassidens*). Each click generally has four to eight cycles and a duration of 15–70 μs. Peak-to-peak source levels can be very high, from 210 to over 225 dB re 1 μPa at 1 m. High-intensity signals from dolphins generally are broadband and can contain frequencies beyond 100 kHz. The frequencies of dolphin clicks vary almost linearly with the signal intensity, such that, as the peak frequency of echolocation signals increases, the intensity of clicks increases (Au and Suthers 2014).

All odontocetes studied thus far produce echolocation signals using one or two pairs of phonic lips located in the nasal passages. The lips contain bursae, which are rod-like fatty structures situated just below the blowhole (AB, PB in Fig. 12.11b). The phonic lips produce both echolocation clicks and communication whistles (Cranford et al. 1996).

Amundin (1991) and Huggenberger et al. (2009) studied click-production in the harbor porpoise, which can serve as a general example for odontocetes other than sperm whales. Figure 12.11 shows an overview and details of the harbor porpoise sound-producing apparatus (Huggenberger et al. 2009). Air passages are shown in blue, fat in yellow, bone in white, and

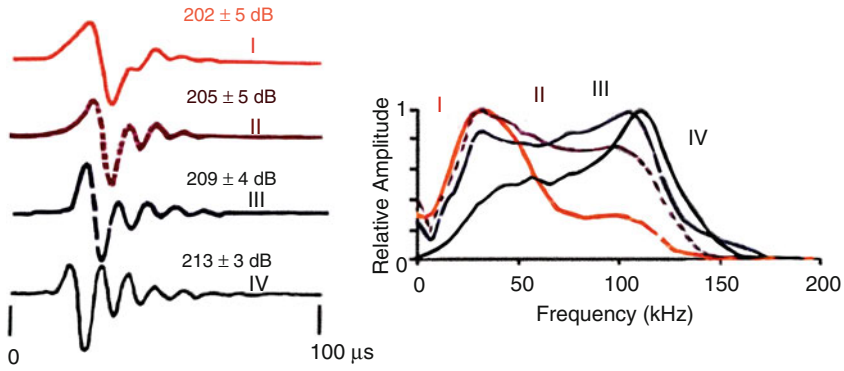



**Fig. 12.10** Left: Waveform of false killer whale biosonar signals with increasing averaged peak-to-peak source level in dB re 1 μPa (relative amplitudes are drawn). Right: Spectra of the corresponding signal type showing increasing peak-frequency with increasing signal amplitude. Adapted by permission from Springer Nature. Au WWL, Suthers RA. Production of Biosonar Signals: Structure and Form, pp. 61–105, in Surlykke A, Nachtigall PE, Fay RR, Popper AN (eds) Biosonar. Springer, New York, NY, USA; https://link.springer.com/chapter/10.1007/978-1-4614-9146-0_3. © Springer Nature, 2014. All rights reserved

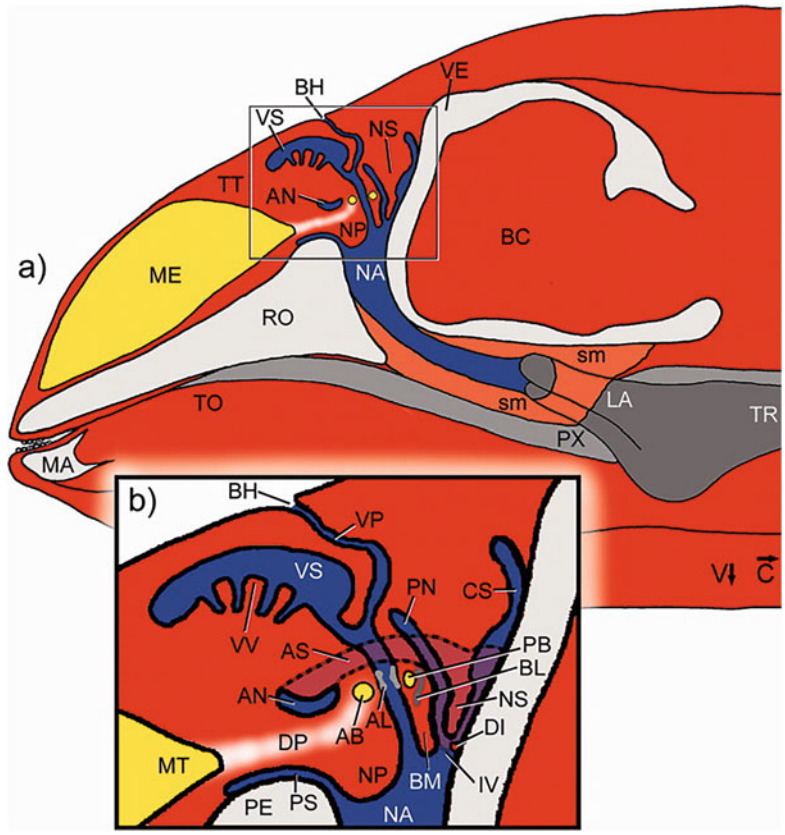



**Fig. 12.11** Schematic sagittal reconstruction of the head of an adult harbor porpoise showing the nasal structures and the position of the larynx (LA). (**a**) Overview. (**b**) Detail of boxed area in (**a**). Blue: air spaces of the upper respiratory tract; gray: digestive system; light gray: cartilage and bone of the skull; yellow: fat bodies. AB: rostral bursa cantantis; AL: rostral phonic lip; AN: anterior nasofrontal sac; AS: angle of nasofrontal sac; BC: brain cavity; BH: blowhole; BL: blowhole ligament; BM: blowhole ligament septum; C: caudal; CS: caudal sac; DI: diagonal membrane; DP: low density pathway; IV: inferior vestibulum; LA: larynx; MA: mandible; ME: melon; MT: melon terminus; NA: nasal passage; NP: nasal plug; NS: nasofrontal septum; PB: caudal bursa cantantis; PE: premaxillary eminence; PN: posterior nasofrontal sac; PS: premaxillary sac; PX: pharynx; RO: rostrum; sm, sphincter muscle of larynx; TO: tongue; TR: trachea; TT: connective tissue theca; V: ventral; VE: vertex of skull; VP: vestibulum of nasal passage; VS: vestibular sac; VV: folded ventral wall of vestibular sac. Reprinted with permission from John Wiley and Sons. Huggenberger S, Rauschmann MA, Vogl TJ, Oelschläger HHA. Functional Morphology of the Nasal Complex in the Harbor Porpoise (*Phocoena phocoena* L.). The Anatomical Record 292:902–920; https://anatomypubs.onlinelibrary.wiley.com/doi/full/10.1002/ar.20854.
© John Wiley and Sons, 2009. All rights reserved

other tissues in red. Air in the bony nares (NA) is pressurized by the nasopharyngeal pouch and the sphincter muscle of the larynx (sm), possibly with help of the piston-like action of the rostral end of the larynx (LA) and epiglottis (Ridgway and Carter 1988). The nasal plug (NP) and the blowhole ligament septum (BM) control the flow of pressurized air past the phonic lip pair (AL: Anterior Lip/PL: Posterior Lip) in each naris resulting in a click-like vibration in the bursae (Anterior Bursa, AB and Posterior Bursa, PB), primarily on the right-side. Each click projects from the bursae through a low-density pathway (DP) to the melon (ME) and from there to the water. This low-density pathway (DP) is characteristic for the families Phocoenidae (porpoises)

and Cephalorhynchinae (small dolphins). In the bottlenose dolphin, and most other delphinids, the anterior bursa (AB) directly abuts the melon. The small amount of air needed to produce a single click ends up in the vestibular air sac (VS) and eventually is re-cycled to the nasal cavity (NA), rather than exhaled through the blow hole (BH) (Norris et al. 1971; Dormer 1979). This process appears to be the same in all odontocetes.

Dormer (1979) showed that in three delphinids, the right pair of phonic lips produces high-frequency clicks, the left pair produces whistles. Whistles, like clicks, are also transmitted to the melon and into the water but are much less directional due to their lower frequencies. There is conflicting evidence for click-production by the left pair of phonic lips (Madsen et al. 2013; Cranford et al. 2011, 2015). Critically designed experiments and field recordings are needed to elucidate the full function of the left pair of phonic lips, particularly in species such as porpoises that do not whistle.

In dolphins, porpoises, and river dolphins, the melon (ME in Fig. 12.11) and associated tissues are the primary structures for transmitting echolocation clicks from the phonic lips to the water (Cranford et al. 1996). In the bottlenose dolphin melon, fat is not homogeneous; rather it is composed of varying amounts of triglycerides and wax esters that differentially affect the sound transmission velocity through the melon (Au 1993, 2015). The same is true for the harbor porpoise (Au et al. 2006; Madsen et al. 2010), where the melon contains mainly triglycerides, probably of many different types (chain lengths and degree of saturation) producing different densities (acoustical impedances). The lowest density is near the low-density pathway (DP in Fig. 12.11), while the highest density approximates that of seawater and occurs in the dorsal part of the melon about four centimeters caudal to the upper lip of the harbor porpoise (Kuroda et al. 2015).

The density of muscle and connective tissue above and lateral to the melon (TT in Fig. 12.11) is greater than the density of the melon tissue and keeps sound from leaking out of the melon. In dolphins and the harbor porpoise, a vestibular air sac (VS) is associated with the melon and also acts like a shield to preventing sound leakage. New results indicate that the melon of the harbor porpoise functions as an acoustic waveguide (Wei et al. 2017, 2018).

The foreheads of beaked whales (Ziphiidae) and the two pygmy sperm whales (family Kogiidae) are quite different. Here, the anterior bursae lie against a spermaceti organ filled with wax esters (Cranford et al. 1996). The spermaceti organ abuts the melon, so an echolocation click first passes through the spermaceti organ into the melon and out into the sea. Beaked whales have an extensive sheet of thick, dense, connective tissue rather than air sacs above the spermaceti organ and melon (Cranford et al. 2008). Beaked whales dive deep and hunt at depths of more than 1000 m (Johnson et al. 2006). At such extreme pressures, air sacs would collapse, but the structural adaptation of the forehead would still protect against acoustic leakage from the melon. Song et al. (2015) measured the acoustical properties of the melon in pygmy sperm whale (*Kogia breviceps*). The density of the melon tissue, and the velocity and impedance of sound are highest in the center of the melon. These physical characteristics keep sound from leaking through connective and muscular tissue surrounding the melon. In addition, air sacs above the spermaceti organ of *Kogia* keep sound in the spermaceti organ. It is unknown how deep *Kogia* dives, but the presence of air sacs above the spermaceti organ suggests that it does not dive as deeply as beaked whales. *Kogia* has extreme right-sided asymmetry of the skull bones, the function of which remains unclear.

The bioacoustical system of the sperm whale differs from all other odontocetes (Cranford et al. 1996). Sperm whales (*Physeter macrocephalus*) have only the right pair of phonic lips, which projects to the tip of the giant rostrum (Fig. 12.12). Click-production is essentially like that of other odontocetes. Air is pressurized in the right naris (Rn) causing a click from the right pair of phonic lips (Mo). A very small amount of sound energy escapes through the distal air sac (Di) at click-production ($P_0$ Fig. 12.12b). The major portion of sound energy projects back

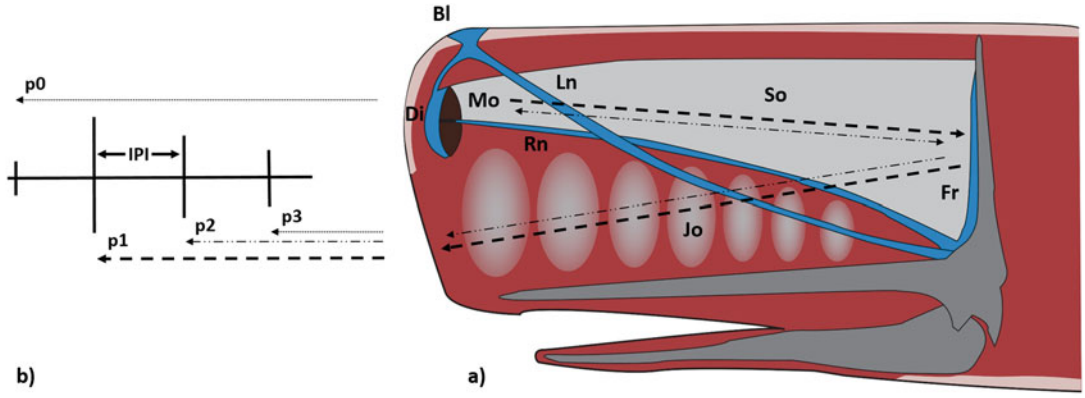



**Fig. 12.12** A schematic drawing of a sperm whale head. *Bl* Blow hole; *Di* Distal air sac; *Fr* Frontal air sac; *Jo* Junk organ; *Ln* Left naris; *Mo* Monkey lips (*museau de singe*); *Rn* Right naris; *So* Spermaceti organ. (**a**) communication or coda clicks and (**b**) echolocation clicks, p1 being the strongest. According to the bent horn model, the production of an intense echolocation click (the solid black dashed lines and p1 in **b**) generates multiple weaker pulses (p2, p3, p4 in **b**) owing to reverberation of the initial sound (p1) between Di and Fr (the thin dashed lines). The whale can modify click generation to produce coda, or weaker communication clicks (the red solid line). This indicates that the whale can somehow control where the click, generated by the monkey lips (Mo), reflects off the frontal air sac (Fr) thus exiting near the distal air sac (Di). Modified from Caruso et al. (2015). © Caruso et al. 2015; https://doi.org/10.1371/journal.pone.0144503. Licensed under CC BY 4.0; https://creativecommons.org/licenses/by/4.0/.

through the spermaceti organ (So, heavy dashed line), hits the frontal air sac (Fr) and is reflected through the "junk" (Jo, heavy dashed line) into the water as a powerful and broadband click ($P_1$ in Fig. 12.12b). The sperm whale $P_1$ click is the most powerful biological sound known (with maximum source levels of 236 dB re 1 µPa rms at 1 m, Møhl et al. 2003), and is probably used as a long-distance biosonar probe signal (see Fig. 12.13b). But it has been proposed that these powerful clicks could stun prey. Norris and Møhl (1983) suggested a "big bang theory" for bottlenose dolphins and sperm whales that produce especially loud, single pulses (or bangs). These pulses could debilitate prey for easy capture, but this has never been proven. In fact, a new study using D-tags on sperm whales recorded no "big bangs," but normal odontocete prey capture behavior (Fais et al. 2016).

A fraction of $P_1$ energy reflects from the distal air sac causing a $P_2$ click to be emitted at a delay consistent with the length of the head (spermaceti organ). The reverberation continues ($P_1$ to $P_4$ in Figs. 12.12b and 12.13a), resulting in a multi-pulse structure. Cranford et al. (1996) proposed that the spermaceti organ and the junk are homologous with the posterior and anterior bursae in the dolphin, respectively.

Although the sound-generating apparatus is basically similar in odontocetes, the outgoing sound from the melon can differ substantially among species. Initially, the action of the phonic lips, controlled by pneumatic pressure, influences the intensity of the click. Stronger hammer-action of a phonic lip pair means the transmission of more intense and higher-frequency clicks (Finneran et al. 2014; Fig. 12.10).

During orientation, most delphinids produce short, broadband echolocation clicks (Au 1993) often of high intensity. They produce less intense, but rapidly repeated clicks, analogous to a bat's buzz when approaching objects or prey (see Fig. 12.1). A single click of a wild white-beaked dolphin lasts about 15 µs and has energy from about 30 kHz to over 200 kHz (Rasmussen and Miller 2002). The sperm whale also fits into this category (Møhl et al. 2003) with a broadband $P_1$ click (Fig. 12.13b).

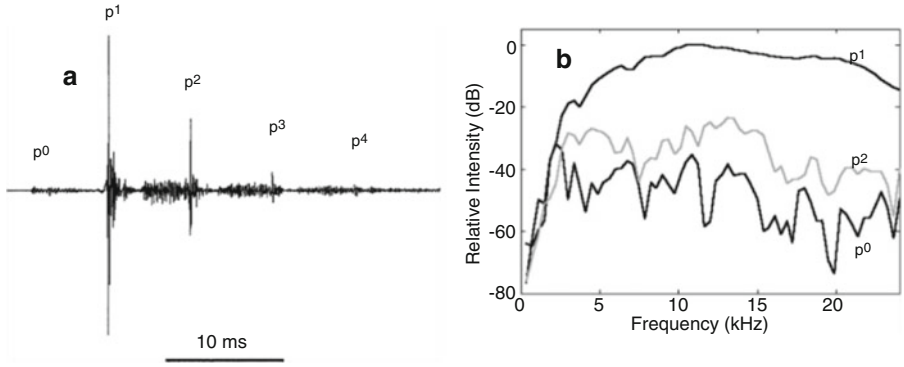



**Fig. 12.13** Multi-pulse structure of a sperm whale click. The $P_1$ click is the most intense and broadest in frequency. It is the most powerful biological sound known. The following clicks of decreasing amplitude ($P_2$–$P_4$) are caused by reverberations in the nose of the whale (see also Fig. 12.12). From Møhl et al. (2003). © Acoustical Society of America, 2003. All rights reserved

At present, it seems that the modulation of clicks in the harbor porpoise occurs in the whale's forehead and that the basic echolocation signals entering the forehead are short-duration, broadband clicks. Madsen et al. (2010) used contact hydrophones to show that a harbor porpoise click recorded near the right (or left) phonic lip pair is broadband. The same click recorded on the melon, along the midline of the animal near the exit point of the sound, has the typical polycyclic narrowband structure. The narrowband high-frequency click (Fig. 12.14) somehow results from the melon and associated tissues, but the details of this mechanism are unknown.

Beaked whales regularly use frequency-modulated up-swept clicks for orientation and when searching for prey. These are relatively broadband and about 200 μs long (Fig. 12.15). Clicks used during prey capture in the buzz are less than 100 μs long, slightly more broadband than the regular clicks and similar to dolphin clicks. It is unknown how the upsweep of the regular click is generated, but by analogy to the porpoise, the basic signal is likely a broadband click somehow shaped in the forehead of the whale.

The directionality of the echolocation sound beam in odontocetes has been studied for many

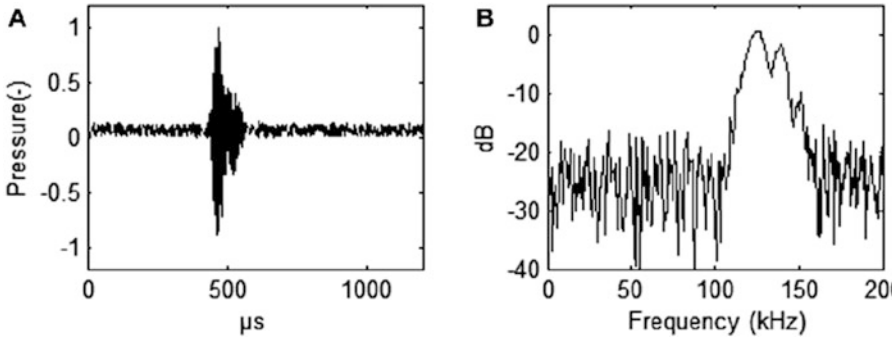



**Fig. 12.14** (**a**) Echolocation click from a harbor porpoise. (**b**) Spectrum of a harbor porpoise click. The harbor porpoise is one of several smaller toothed whales that use a high-frequency narrowband echolocation click (Galatius et al. 2019). From Fig. 12.1 in Miller and Wahlberg (2013); © Miller and Wahlberg 2013; https://doi.org/10.3389/fphys.2013.00052. Licenced under CC BY 3.0; https://creativecommons.org/licenses/by/3.0/

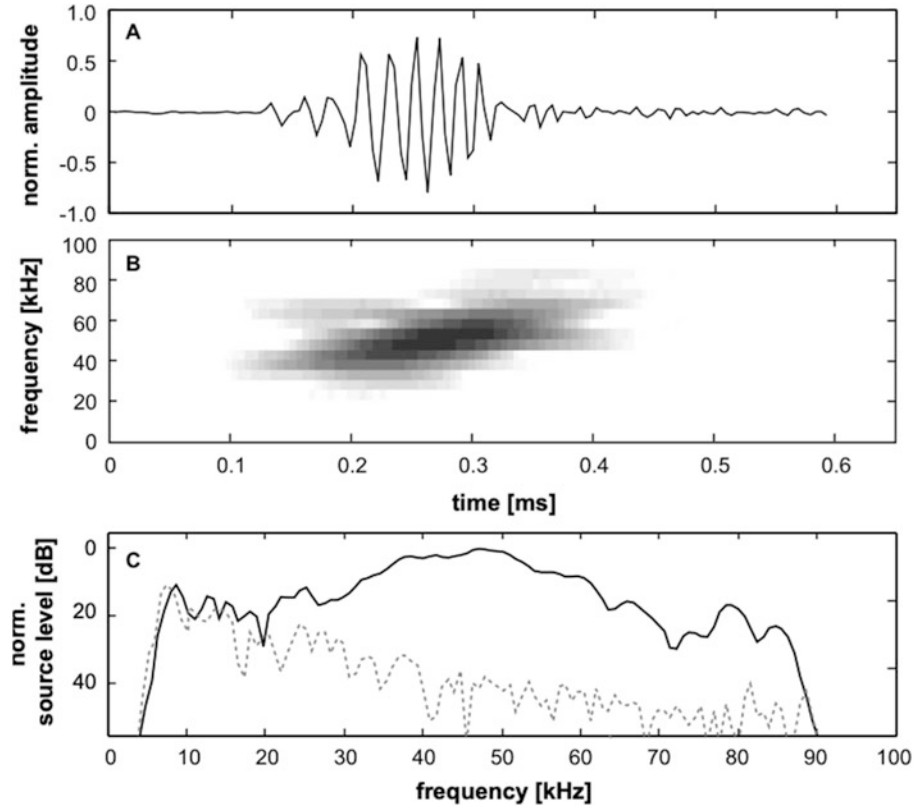



**Fig. 12.15** Beaked whale click waveform (**a**), spectrogram (**b** Hann window, 40-point FFT, 98% overlap), and spectrum (**c** Hann window, 256-point FFT; dashed line shows ambient noise). Baumann-Pickering et al. (2010). © Acoustical Society of America, 2010. All rights reserved

years (Au 1993, 2015; Au et al. 1985, 1986, 1999; Kloepper et al. 2012; Koblitz et al. 2012). Recent work reveals that odontocetes control the shape and direction of the beam (Moore et al. 2008; Wisniewska et al. 2015). A bottlenose dolphin with its head stationary and its mouth on a biteplate moved its sound beam by 26° to the left and 21° to the right when echolocating a movable sphere 9 m away (Moore et al. 2008). Wisniewska et al. (2015) used two-dimensional hydrophone arrays to verify that harbor porpoises approaching a target (a dead fish) voluntarily change the diameter of their echolocation beam to increase the ensonified area by 100–200%, while reducing the interval between clicks in the buzz phase just before prey capture (Fig. 12.16). These changes are analogous to what a bat will do when capturing an insect (Jakobsen et al. 2015). Wild Amazon river dolphins (*Inia geoffrensis*) also increase the beam width during prey capture (Ladegaard et al. 2017). Increasing the beam width helps the porpoise (or bat) track a moving prey at close proximity. Presumably, the musculature around the melon helps control the beam width and direction in porpoises and dolphins (Moore et al. 2008), but this needs verification.

The direction of the sound beam from the head of a porpoise carcass can be changed by artificially inflating the vestibular air sacs (Miller 2010). With no air in the vestibular air sacs, a broadband click generated by a small hydrophone between the right pair of phonic lips projects left of the midline and vice versa with an artificial click generated between the left phonic lip pair. With air in the vestibular air sacs, the artificial clicks project out the midline (Fig. 12.17; see also

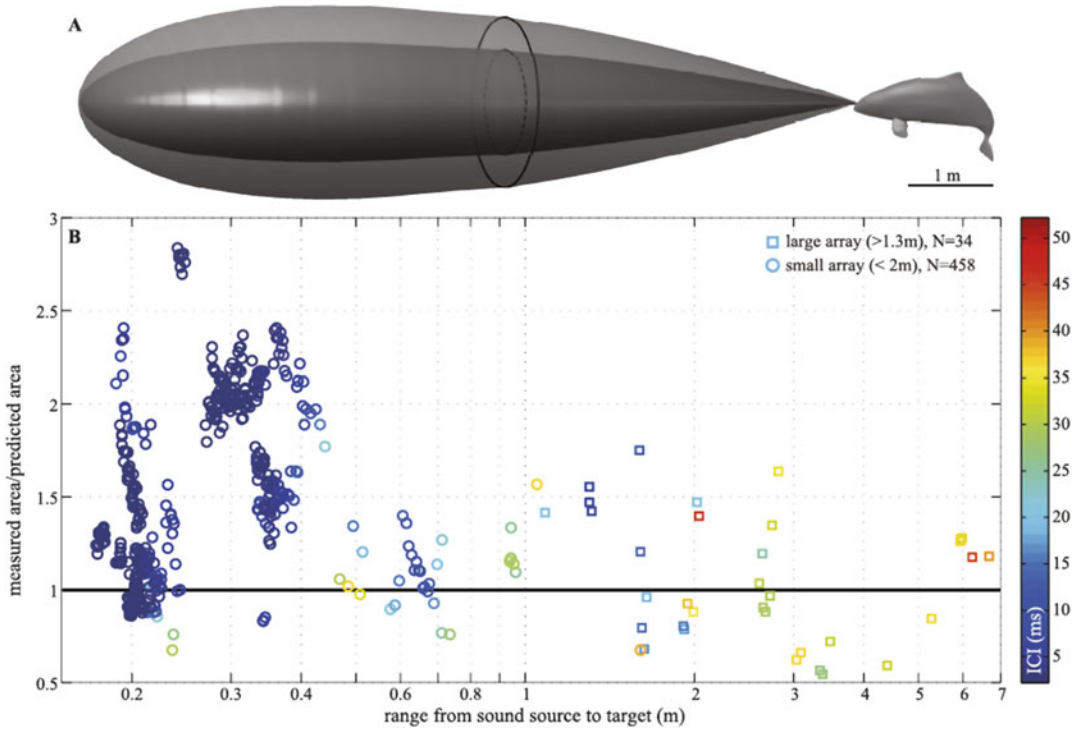



**Fig. 12.16** The harbor porpoise can increase the ensonified area by nearly 200% during the buzz phase with short inter-click intervals (ICI in **b**, blue). The large diameter circle (solid in **a**) illustrates the beam width for clicks with short intervals. The small diameter circle (dashed in **a**) shows the beam width of clicks with longer intervals emitted in the search phase at longer distances (ICI in **b**, red). © Wisniewska et al. 2015; https://elifesciences.org/articles/05651. Licensed under CC BY 4.0; https://creativecommons.org/licenses/by/4.0/. All rights reserved

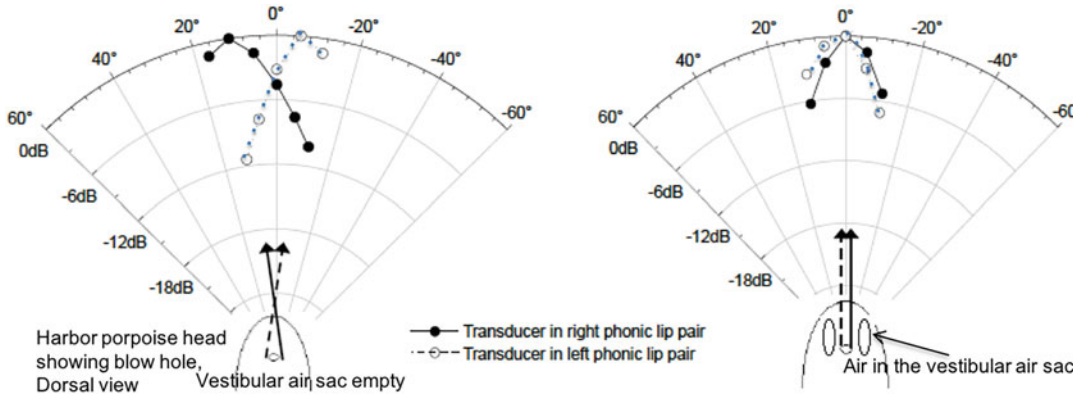



**Fig. 12.17** Short broadband artificial clicks generated between the phonic lips (right lip: solid arrow and curve; left lip: dashed arrow and curve) of a cadaver harbor porpoise. With air in the vestibular air sacs (right image), the clicks emerge at the midline. Without air in the vestibular air sacs (left image), the clicks emerge on either side of the midline depending on where the artificial click was generated (clicks generated between the right pair of phonic lips emerge to the left and vice versa). Adapted with permission from Miller LA (2010); Prey Capture by Harbor Porpoises (*Phocoena phocoena*): A Comparison Between Echolocators in the Field and in Captivity; J Marine Acoust Soc Jpn 37 (3):156–168. © The Marine Acoustics Society of Japan, 2010

Starkhammar et al. 2011; Cranford et al. 2014). Incidentally, the exiting click remained broadband in these experiments indicating that the living melon and associated tissues are necessary for producing a high-frequency, narrowband click typical for the harbor porpoise (Madsen et al. 2010).

The primordial odontocete echolocation signal was probably a short, broadband click similar to the clicks used by most living dolphins and the sperm whale (Fig. 12.10, left). In contrast, the La Plata dolphin (*Pontoporia blainvillei*), six small dolphins (family Delphinidae), all porpoises (family Phocoenidae, six species with four documented), and the pygmy and dwarf sperm whales (family Kogiidae) use narrowband, high-frequency (NBHF) echolocation clicks (see Fig. 12.14). The change from broadband to NBHF echolocation clicks could reflect predation pressure by killer whales (and their ancestors), as well as environmental factors (Andersen and Amundin 1976; Madsen et al. 2005; Morisaka and Connor 2007; Miller and Wahlberg 2013; Galatius et al. 2019). NBHF clicks appear to be generated in the melon and associated tissues (Madsen et al. 2010). It is assumed that all odontocetes can control the amplitude of echolocation clicks, steer the sound beam, and manipulate its width (Moore et al. 2008; Wisniewska et al. 2015). These features are of obvious advantage for detecting and tracking prey. There are rich possibilities in future research of sound production and the use of echolocation by odontocete whales.

### 12.5.2 Hearing Anatomy and Echolocation Abilities

We refer to Vol. 2 Chap. 9 on aquatic mammals for more detail on hearing anatomy and abilities. Here, we focus on the hearing abilities of odontocetes as they relate to the tasks of obstacle and prey detection by echolocation.

Experimental studies show that the bottlenose dolphin (Li et al. 2011), the false killer whale (Nachtigall and Supin 2008), and the harbor porpoise (Linnenschmidt et al. 2012, 2013) have voluntary control over the level of the emitted click and of their auditory sensitivity during echolocation tasks. The results from the harbor porpoise clearly illustrate active hearing during the echolocation of targets: the porpoise maintains a constant level of auditory perception independent of target distance. If the distance to a target is doubled, the level of a click impinging on the target is halved (−6 dB). To compensate for this, the porpoise doubles the level of the outgoing click (+6 dB), keeping the level of the incident sound on the target constant and independent of distance (within a certain range). However, the returning echo is halved (−6 dB) at double the distance. Linnenschmidt et al. (2012) showed that there is an "automatic gain control" in the auditory system of the porpoise such that its hearing increases in sensitivity by about +6 dB to compensate for the loss in the echo level over double the distance. Without compensating for the level of the outgoing click and the gain control in the auditory system, the echo level would drop by 1/4 (−12 dB) per doubling of distance to the target, making echolocation more difficult for the whale.

Toothed whales obviously find their prey using echolocation, but how they discriminate between prey species is not known and, to our knowledge, has not been studied experimentally. Probably the most spectacular use of echolocation to find prey is shown by bottlenose dolphins in the Grand Bahamas. The dolphins often find fish under the sand using their echolocation and stick their proboscis down in the sand, sometimes to the pectoral fins, and come up with a fish in their mouths (Rossbach and Herzing 1997). What echo information they use for this unusual behavior is unknown. Harbor porpoises can discriminate between identical spheres of different materials (Wisniewska et al. 2012). Three harbor porpoises were easily able to distinguish between an aluminum sphere and spheres of plexiglas, PVC, and brass. Two of the three had problems differentiating aluminum from steel spheres. The spectra of these two spheres were very similar, so we assume the harbor porpoises were using spectral information to detect the differences among

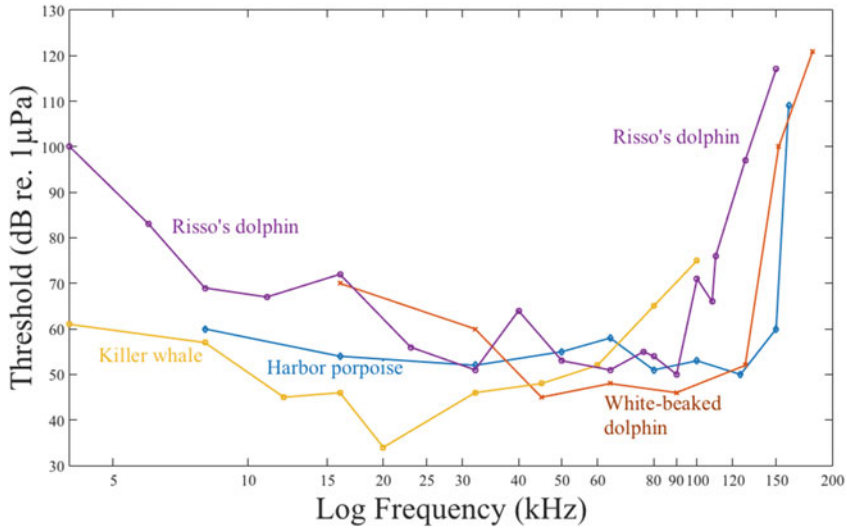



**Fig. 12.18** Underwater audiograms of four odontocetes. Blue: Harbor porpoise behavioral audiogram using a 50-ms sound stimulus (Kastelein et al. 2010). Orange: White-beaked dolphin auditory evoked response audiogram using a 1-s sinusoidal amplitude-modulated stimulus (Nachtigall et al. 2008). Purple: Risso's dolphin (*Grampus griseus*) auditory evoked response audiogram using a 20-ms sinusoidal amplitude-modulated stimulus (Nachtigall et al. 2005). Yellow: Killer whale average behavioral audiogram of two animals using a 2-s tone (Szymanski et al. 1999)

the spheres. Perhaps they also use spectral information together with target strength to distinguish between different fish species.

All echolocating toothed whales have a U-shaped audiogram (Fig. 12.18) and a broad range of hearing extending up to 200 kHz. In general, the hearing of odontocetes is most sensitive at the frequencies used for echolocation. For example, the harbor porpoise, a narrow-band high-frequency species, is most sensitive at around 130 kHz, the peak frequency of its narrow band signal. The killer whale uses lower frequencies in its echolocation signals and its best hearing is accordingly lower (Fig. 12.18).

## 12.6 Echolocation in Birds

The oilbird (*Steatornis caripensis,* family Steatornithidae), and a subset of the swiftlets, family Apodidae (about 16 of 27 species, currently including *Aerodramus spp* and *Collocalia troglodytes*) are the only birds known to echolocate (Griffin 1958; Novick 1959; Chantler et al. 1999; Price et al. 2004). Neither seem to use echolocation to find food, but rather for crude orientation in dark caves or tunnels where they roost and nest. Arguably, bird echolocation systems are not a highly evolved sensory specialization in the same sense as in bats and odontocetes.

Disregarding nesting habits, oilbirds and swiftlets have very different ecologies. Oilbirds are nocturnal fruit-eaters from the tropical part of South America (Chantler et al. 1999). Swiftlets occur across the Indo-Pacific and use vision to locate insect prey during the day. There are records of swiftlets hunting at dusk, but it is unclear if they use echolocation during this activity (Price et al. 2004; Fullard et al. 1993).

### 12.6.1 Sound Production and Signal Characteristics

Like other birds, oilbirds and swiftlets produce sounds, including their biosonar signals, by inducing vibrations in air passed by membranous structures in their syrinx (see Vol. 2, Chap. 6).

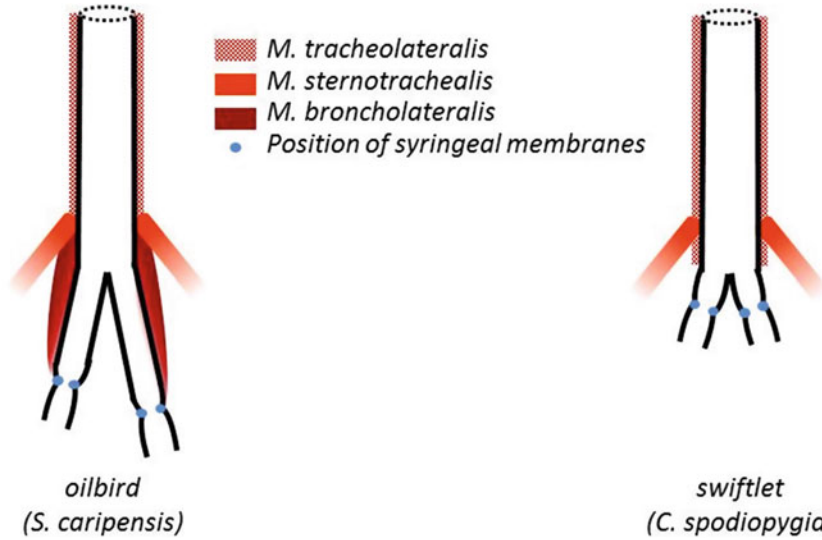



**Fig. 12.19** Schematic of syrinx anatomy in the oilbird (based on Suthers and Hector 1988, Fig. 12.2) and the Australian grey swiftlet (*Aerodramus* (formerly *Collocalia*) *spodiopygia*; based on Suthers and Hector 1982, Fig. 12.2), showing the trachea and its bifurcation into the two bronchi. Note the lack of intrinsic syringeal muscles (*mm. broncholateralis*) in the swiftlet. Note also the asymmetry of the bronchial oilbird syrinx with a more cranial placement of the right semi-syrinx. Adapted by S. Brinkløv

Suthers and Hector (1982, 1985) revealed distinct differences in the syringeal morphology of oilbirds and swiftlets (Fig. 12.19) but proposed similar sound production mechanisms in both. Oilbirds have a bronchial syrinx located caudal to the tracheal bifurcation. The two half-syringes are placed with bilateral asymmetry in the two bronchi (Suthers and Hector 1985). The swiftlet syrinx is tracheobronchial (i.e., located where the trachea splits into the two bronchi; Suthers and Hector 1982).

Suthers and Hector suggested that biosonar signals in both oilbirds and swiftlets are produced as a contraction of the extrinsic sternotrachealis muscles pulls the trachea caudal. This reduces tension across the syrinx and causes the syringeal membranes to fold into the syrinx lumen, where they induce vibrations of the expiratory airflow. Contrary to their other vocalizations, oilbirds and swiftlets actively terminate their echolocation clicks but do so by using different sets of muscles. In oilbirds, termination is controlled by contraction of the broncholateralis muscles intrinsic to the syrinx (Suthers and Hector 1985). Swiftlets lack intrinsic syringeal muscles (Fig. 12.19) and instead contract extrinsic tracheolateralis muscles to terminate their echolocation clicks (Suthers and Hector 1982).

Bird biosonar signals are relatively broadband and without structured frequency changes over time (Pye 1980). In this sense, they resemble the tongue-clicks of rousettes bats more than the signals produced by other echolocators, but with a narrower frequency range, longer duration, and lacking similarly well-defined on- and offsets (Fig. 12.20).

In the wild, oilbirds emit click-bursts of two or more single clicks in rapid succession (Fig. 12.20). Their clicks and click intervals are stereotyped within such a burst, with click durations of 0.5–1 ms and click intervals of ~2.5 ms. Clicks recorded from oilbirds in the wild have the most energy around 10–15 kHz but extend from 7 to 23 kHz measured at −6 dB from the peak frequency (Brinkløv et al. 2017). The intervals between click-bursts are more variable, but often around 200 ms (Griffin 1953). Each click-burst is perceived by human ears as

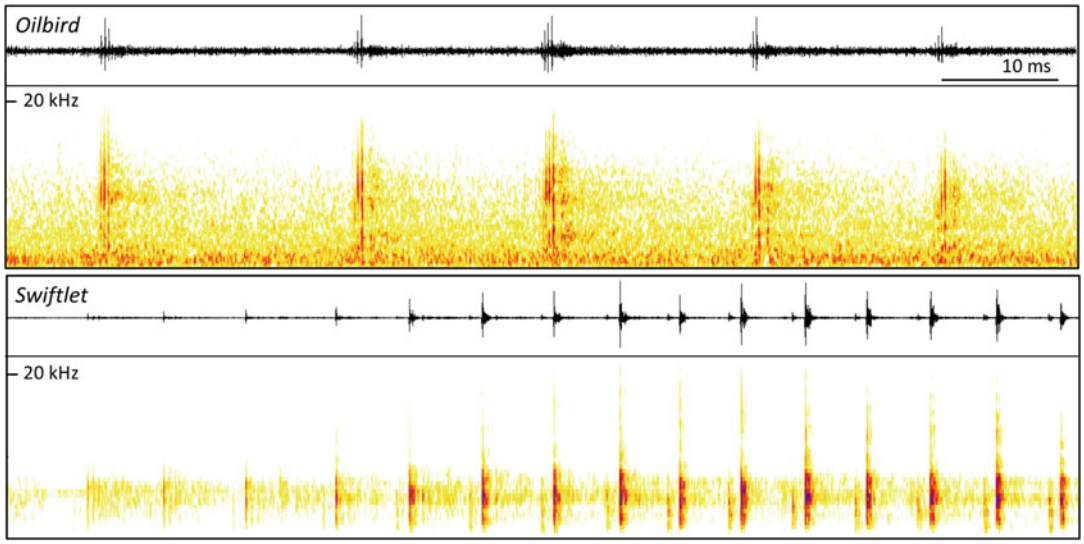



**Fig. 12.20** Waveform and spectrogram displays of bird echolocation click sequences. Top panel: oilbird (*Steatornis caripensis*) exiting cave roost, recorded at Dunstan's Cave, Asa Wright Nature Centre, Trinidad. Bottom panel: swiftlet (*Aerodramus unicolor*) returning to its nest in a Sri Lankan railway tunnel. The overall timescale is 1 s, frequency scale is from 0 to 20 kHz. Spectrogram settings: FFT size 256, Hann window, 98% overlap. Both recordings are high-pass filtered at 1 kHz (second order Butterworth filter)

one coherent sound (Konishi and Knudsen 1979). It is unresolved whether the number of individual clicks in a burst has functional meaning to the oilbird, but recent studies indicate that oilbirds may add click subunits to a burst as a means to increase overall burst energy and, as a result, the echolocation range (Brinkløv et al. 2017). Click-bursts typically have source levels of around 100 dB re 20 μPa rms at 1 m (Brinkløv et al. 2017).

Data from captive oilbirds differ somewhat from field recordings. Konishi and Knudsen (1979) reported that oilbird signals had most energy around 2 kHz and described each click as a pulse-like sound burst of 20 ms or more. Suthers and Hector (1985) described a large signal variation including continuous pulsed signals of 40–80 ms and shorter single or double pulses. This difference between field and captive data possibly indicates that the sounds of captive birds do not accurately reflect the echolocation behavior of birds in the wild since vocalization could be affected by reverberant confines or the stress of handling/being restrained.

Swiftlets emit biosonar signals either as single or double clicks (two single clicks in rapid succession, Thomassen et al. 2004; Fig. 12.20). As in oilbirds, it is unclear if the difference between single and double clicks has functional meaning to the swiftlets or is merely an artifact of the sound production mechanism (Suthers and Hector 1982). Of 12 swiftlet species studied, only the Atui swiftlet (*Aerodramus sawtelli*) appears to consistently produce single clicks (Fullard et al. 1993), while the rest emit both single and, more often, double-clicks. Each click of a pair is 1–8 ms long, with the second often of higher amplitude and slightly longer duration (Griffin and Suthers 1970; Suthers and Hector 1982; Coles et al. 1987). Clicks within a pair have intervals of 1–25 ms and click-pairs are emitted at intervals of 50–350 ms. Swiftlet clicks have most energy below 10 kHz (see spectrogram in Fig. 12.20).

### 12.6.2 Hearing Anatomy and Echolocation Abilities

While the auditory systems of echolocating bats and odontocetes include specializations that confer increased acuity and sensitivity, only a few such morphological or neurological specializations have been found in echolocating birds. Tomassen et al. (2007) used three-dimensional, micro-CT scans to model the middle ear function of a range of swiftlet species. They found no morphological adaptations in the middle ear single bone-lever system of the birds (Fig. 12.21) to improve impedance-matching in echolocating compared to non-echolocating species. Both had low tympanum-to-oval-window ratios relative to bird auditory specialists such as owls. Birds have a straight, rather than coiled cochlea (Fig. 12.21) and generally do not hear much above 10 kHz (Fig. 12.9, also see Manley 1990, p. 238).

While peripheral auditory adaptations for echolocation seem absent in birds, there is some evidence that certain of the brain nuclei involved in auditory processing are enlarged in echolocating bird species. Thomassen (2005) found that echolocating swiftlets have larger nuclei magnocellularis and nuclei laminaris compared to non-echolocating swiftlets, structures that are both involved in temporal coding of auditory stimuli. The nucleus angularis appears to be enlarged in oilbirds (Kubke et al. 2004) and is known to process intensity information in barn owls (*Tyto alba*). Iwaniuk et al. (2006) concluded that oilbirds and swiftlets may have enlarged MLds (nucleus mesencephalicus lateralis, pars dorsalis), a structure homologous to the mammalian inferior colliculus. However, this enlargement was only apparent compared to closely related non-echolocating species, not to non-echolocating birds in general.

The hearing abilities of both oilbirds and swiftlets have been tested using neurophysiological approaches and indirectly through obstacle avoidance experiments. Measurements of cochlear and evoked potentials from the forebrain nucleus of anesthetized oilbirds empirically support the absence of inner ear specializations for echolocation. Oilbirds appear to be more or less insensitive to frequencies above 6 kHz and their best auditory sensitivity is at ~2 kHz (Fig. 12.9, and Konishi and Knudsen 1979). Single neuron recordings from the midbrain auditory nucleus of the echolocating Australian grey swiftlet showed

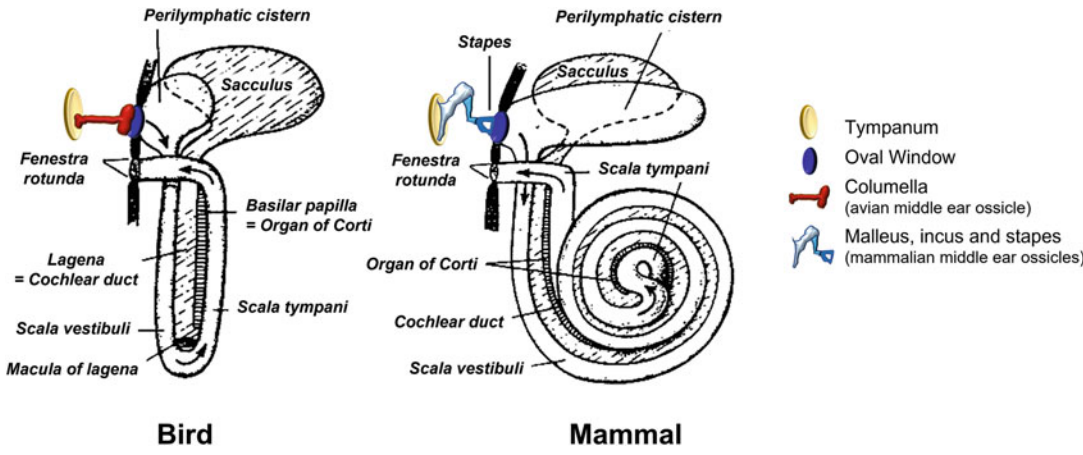



**Fig. 12.21** Overview of avian and mammalian middle and inner ear anatomy. Left: Birds have a single middle ear bone (columella) and a straight cochlea. Right: Mammals have three middle ear bones (malleus, incus, and stapes) and a coiled cochlea. Adapted by permission from Springer Nature. Manley GA, Peripheral hearing mechanisms in reptiles and birds; https://www.springer.com/gp/book/9783642836176. © Springer Nature, 1990. All rights reserved

best thresholds at 1–5 kHz (Fig. 12.9 and Coles et al. 1987). Hence, both oilbirds and swiftlets appear to have the 'standard' bird hearing range, with lowest thresholds between 2 and 4 kHz and poor sensitivity above 10 kHz (Dooling 1980). Curiously, it appears that oilbirds in the wild emit echolocation clicks that are not well-aligned to their best area of hearing. The lack of external ear structures in oilbirds and swiftlets means that directional cues occur at frequencies predicted by head size.

With echolocation signals matching their most sensitive area of hearing, oilbirds and swiftlets should detect objects down to at least 17 cm in diameter, equal to the wavelength of the signal at 2 kHz. For Oilbirds, this prediction is supported by obstacle-avoidance experiments, suggesting that they detect discs 20 cm in diameter suspended from the ceiling of their cave roost (Konishi and Knudsen 1979). However, detection thresholds between 0.6 and 2 cm have been found for swiftlets (Griffin and Suthers 1970; Fenton 1975; Griffin and Thompson 1982; Smyth and Roberts 1983), indicating that they may somehow extract echo information from the upper, albeit weaker, frequency range of their signals.

Like bats and odontocetes, oilbirds and swiftlets detect obstacles in dark spaces using echolocation. Unlike bats and odontocetes, echolocating birds, even the nocturnal oilbird, are also vision specialists and presumably do not forage by echolocation. The importance of vision in oilbirds is reflected in their specialized retinal morphology with multiple layers of photoreceptors (Martin et al. 2004). Initial behavioral experiments revealed that oilbirds flying in darkness consistently produced sounds but could not avoid obstacles if their ears were blocked. With the lights on, the birds, in contrast, produced fewer or no sounds and negotiated obstacles also with their ears blocked (Griffin 1953).

Biosonar signals of birds are generally stereotyped (Thomassen and Povel 2006) and there is no indication that birds have similar adaptive control over signal frequency as most echolocating bats. However, Brinkløv et al. (2017) recently found that the intensity of oilbird echolocation signals increased on darker nights relative to nights with more ambient light. The higher intensity of click-bursts emitted on darker nights resulted both from an increase in the amplitude of individual clicks and an increase in the number of individual clicks per click-burst. Several studies have noted that swiftlets increase click repetition rate as they approach obstacles (Griffin and Suthers 1970; Coles et al. 1987) and Atiu swiftlets emit signals at higher repetition rate when they enter than when they emerge from their cave roost (Fullard et al. 1993).

Nesting in dark places, such as caves, mines, tunnels, and other places where the lighting is uncertain, is a common feature of the ecology of oilbirds and echolocating swiftlets. Both start clicking as they cross a threshold from light to dark (Fenton 1975; Thomassen 2005; Brinkløv et al. 2017). Neither have been shown to use echolocation for foraging, although oilbirds may be able to detect some of the larger fruits they eat (palm fruits up to 6 cm) by echolocation (Snow 1961, 1962; Bosque et al. 1995).

## 12.7 Orientation and Echolocation in Insectivores and Rodents

### 12.7.1 Echo-Based Orientation in Insectivores: Tenrecs and Shrews

Tenrecs and shrews are small insectivorous mammals that forage in dense vegetation or under leaf-litter (Fig. 12.22). Tenrecs are largely endemic to Madagascar, but shrews have a wide distribution across Eurasia and North America. Both have tiny eyes and a presumably well-developed olfactory sense and emit a variety of sounds. The use of sounds by shrews and tenrecs, as they approach and explore unfamiliar objects in their surroundings, led to initial suggestions that they may use echolocation. However, few studies have successfully tested this hypothesis directly. The current consensus is that shrews and tenrecs may use a simple echo-based orientation system to obtain rough acoustic input about their surroundings at short range beyond their snout and vibrissae. As stated by Siemers et al.



**Fig. 12.22** Photographs (from left) of lowland streaked tenrec (*Hemicentetes semispinosus*), lesser hedgehog tenrec (*Echinops telfairi*), and northern short-tailed shrew (*Blarina brevicauda*). Photo of lowland streaked tenrec by Frank Vassen, 2010, https://commons.wikimedia.org/wiki/File:Lowland_Streaked_Tenrec,_Mantadia,_Madagascar.jpg#filelinks. Photo of lesser hedgehog tenrec by Wilfried Berns, 2006, https://en.wikipedia.org/wiki/Lesser_hedgehog_tenrec#/media/File:Kleiner-igeltanrek-a.jpg. Photo of northern short-tailed shrew by Giles Gonthier, 2007, https://en.wikipedia.org/wiki/Northern_short-tailed_shrew#/media/File:Blarina_brevicauda.jpg. All photos licensed under CC BY 2.0; https://creativecommons.org/licenses/by/2.0/deed.en

(2009): "Except for large and thus strongly reflecting objects, such as a big stone or tree trunk, shrews probably are not able to disentangle echo scenes, but rather derive information on habitat type from the overall call reverberations. This might be comparable to human hearing whether one calls into a forest or into a reverberant cave."

Gould et al. (1964) and Gould (1965) provided the most direct evidence for echo-based orientation in several species of shrews and tenrecs. After unsuccessful attempts to use an obstacle-avoidance set-up, the animals were instead tested using a so-called disc-platform apparatus. They were trained to find and jump onto a platform suspended at a vertical distance below a disc with an area of partial overlap. The location of the overlap was varied at random between trials. Both tenrecs and shrews emitted sounds during this task in the dark, but animals with their ears blocked were less successful in finding and landing on the platform than control animals. The control experiments included two tenrecs that were blindfolded.

Gould (1965) recorded the sound pulses emitted by captive tenrecs (*Echinops telfairi, Hemicentetes semispinosus*, and *Nesogale* (formerly *Microgale*) *dobsoni*) as they explored the disk-platform apparatus. The tenrecs emitted series of tongue clicks, each less than 2 ms long with most energy between 10 and 16 kHz. The clicks were produced as singles, doubles, or in triplets. Streaked tenrecs (*Hemicentetes semispinosus*) emitted clicks of low intensity; while those of *Nesogale dobsoni* were audible to humans at 7 m.

Gould et al. (1964) found that, contrary to the audible pulses of tenrecs, shrews (*Sorex vagrans, S. cinereus, S. palustris, and Blarina brevicauda*) searching for the platform emitted ultrasonic pulses with most energy between 30 and 60 kHz. The pulses were about 5 ms in duration with inter-pulse intervals of about 20 ms. Sanchez et al. (2019) recorded five *Sorex unguiculatus* in three different experimental setups, including soft and hard barrier obstacles. Under all three conditions, the shrews emitted a variety of calls, including clicks and several tonal pulse types ranging in frequency between 5 and 45 kHz with durations of 3–40 ms. While several studies have shown that shrews and tenrecs do show context-dependent changes in vocalization rate, there is little direct evidence for echolocation by these animals (Buchler 1976; Tomasi 1979; Forsman and Malmquist 1988; Siemers et al. 2009; Sanchez et al. 2019).

No morphological adaptations for echolocation have been found in the auditory systems of

tenrecs or shrews. The limited data on hearing in these animals indicate that at least tenrecs hear well across the frequency range of their tongue-clicks. Sales and Pye (1974) reported that the hearing of streaked tenrecs is most sensitive from 2 to 60 kHz. Drexl et al. (2003) used otoacoustic emissions and auditory evoked potentials from the inferior colliculus and the auditory cortex to determine that the auditory range of lesser hedgehog tenrecs (*Echinops telfairi*) extends from 5–50 kHz at 40 dB SPL, with a lowest threshold at 16 kHz. Siemers et al. (2009) report a best hearing range of shrews between 2 and 20 kHz.

### 12.7.2 Echolocation in Rodents

One important test for echolocation is to blind the echolocator. This was done by Griffin (1958) for bats and by Norris et al. (1961) for dolphins. Although such a "blinding test" was not performed, a multifaceted study by He et al. (2021) convincingly suggests soft-furred tree mice (*Typhlomys*) must be added to the list of echolocating animals. Through behavioral experiments in total darkness, filmed with an infrared video camera, they showed that all four species of soft-furred tree mouse emitted acoustic pulses at higher rate and grouped pulses more in complex space than open space and during obstacle avoidance. Further, three species (*T. cinereus*, *T. daloushanensis*, and *T. nanus*) were tested in a disk-platform setup similar to that used by Gould et al. (1964) for shrews and tenrecs. The tree mice spent increased time emitting higher pulse rates on the sector of the disk above the platform before dropping down onto the platform. This preference was lost when their ears were blocked but regained when the ears were unplugged or fitted with hollow tubes. The study also used laboratory house mice (*Mus musculus*) as a control to demonstrate absence of any location preference or sound emission during the disk-platform test. Myriad tests and field studies document the functional use of echolocation by bats and toothed whales, but such studies are not available for insectivores and rodents.

Supplementing the behavioral part of their study, He et al. (2021) also conducted anatomical scans to reveal that the stylohyal bone of soft-furred tree mice is fused with the tympanic bone, which is characteristic of echolocating bats. Lastly, they used genetic analyses to document a strong convergence of hearing-related genes with those of other echolocating mammal groups, including the prestin gene associated with echolocation in bats and toothed whales (Liu et al. 2014). All four species of soft-furred tree mice emit similar short (~2 ms) ultrasonic pulses ranging from 65 to 140 kHz (He et al. 2021).

## 12.8 Are Echolocation Signals also Used for Communication?

Studies on the role of echolocation signals for intraspecific communication have included observations and recordings, playback experiments, and combinations of these approaches. Echolocation signals elicited territorial behavior in foraging spotted bats, served in individual recognition, and assisted in maintaining group adhesion among foraging molossids (Fenton 1995). Furthermore, bats use buzzes (high pulse repetition rates) not only when attacking prey, but also during landing, drinking and by several species in social settings (e.g., Schwartz et al. 2007). Many bat species roost in large groups in caves and emerge at dusk as a group to forage. Several toothed whale species forage in large numbers. Echolocation in bats and odontocetes likely plays a role in maintaining spacing among group members during foraging or during large group movements. However, there has been little research on whether all or only specific animals echolocate while foraging as a group. The benefits of eavesdropping on each other's echolocation signals need to be studied. Groups of flying bats and swimming toothed whales surely eavesdrop on each other's echolocation signals to gain general information about prey location. The energetic cost of sound production for flying bats and for clicking dolphins is negligible (Speakman and Racey 1991; Noren et al. 2017).

Evidence suggests that toothed whales use their echolocation clicks as communication signals. These comprise repeated patterns of rising, falling, or constant click repetition rates up to near 1000 clicks/s. Clicks used for communication by dolphins and porpoises have the same spectral properties as those used for echolocation, but this does not hold true for the coda-clicks of sperm whales, as explained below.

In toothed whales, most is known about the communication role of echolocation clicks from studies of captive harbor porpoises, captive bottlenose dolphins, and wild sperm whales. Porpoises and dolphins communicate with changing click repetition rates, rather like Morse code, without changing the temporal and spectral properties of the clicks (Rasmussen and Miller 2002; Clausen et al. 2010). These "pulse-bursts" (or burst-pulse sounds) of high repetition rate clicks with narrow sound beams are especially good for close range and directed communication (Clausen et al. 2010).

Figure 12.23 shows click rates used in five behavioral contexts between a mother harbor porpoise and her calf. The porpoises used the highest click rates in aggressive encounters, the lowest in grooming and echelon swimming (Clausen et al. 2010). The mother may be aggressive toward her calf and toward males. Aggressive signals were usually higher in intensity and repetition rates and always resulted in the other animal moving away from the emitter. Both mother and calf emitted approach signals, but only the calf emitted contact signals and only the mother emitted grooming signals. Wild harbor porpoises also use rapid click rates for communication (Sørensen et al. 2018).

Bottlenose dolphins use both echolocation clicks and whistles as communication signals. Blomkvist and Amundin (2004) studied two captive female bottlenose dolphins that used high-frequency, high repetition rate pulse-bursts during aggressive behavior. The pulse-bursts lasted up to 900 ms with click repetition rates from 100 to 940 clicks/s. Like the echolocation clicks used for orientation and foraging, the pulses were between 60 and 150 kHz. The metabolic rate of dolphins producing clicks was only slightly greater than that of silent dolphins indicating that echolocation is not energetically costly (Noren et al. 2017).

Several free-ranging species of dolphins (*Tursiops truncatus, Stenella attenuata, S. longirostris, S. frontalis, Orcinus orca,* and *Cephalorhynchus hectori*) use pulse-bursts mostly during affiliative and aggressive behavior (Dawson 1991; Herzing 2000; Lammers et al. 2004). Rasmussen et al. (2016) played back artificial pulse-burst signals (repeated at 300 clicks/s for 2 s) to 21 free-ranging white-beaked dolphins. Rather than responding with aggressive behavior, the dolphins showed mostly a change in swimming direction and swam around the projection equipment, mirroring the retreat of individual captive harbor porpoises receiving an 'aggressive' pulse-burst. The pulse-bursts, or rasps, of Blainville's beaked whale are only emitted at depths below 200 m and composed of a series of short, FM clicks similar to its FM echolocation clicks, except with a lower peak-frequency. The communication context is not known (Arranz et al. 2011).

Sperm whales are social and form social units in subtropical and tropical waters worldwide. Up to 12 females with young of both sexes gather in long-term stable social units. Sperm whales in all ocean basins communicate using rhythmic "coda" clicks (see Fig. 12.12), which are a unique specialization among toothed whales (Watkins and Schevill 1977) and may even signify individual identity. The composition of codas can have many repetitive patterns, such as one click + a group of three clicks: 1 + 3, or 2 + 1 + 1 + 1, 1 + 1 + 3, etc. The coda patterns are not stereotyped; click intervals within a coda can vary and seem to contain information for the receiver. One stable social unit of five adult females, a juvenile male, and a calf in the waters off Dominica used 15 different codas. All individuals in the unit used several codas and one individual used 11 of the 15 codas (Antunes et al. 2011). A recent study (Oliveira et al. 2016) confirmed and extended those of Antunes et al. (2011). Using digital data acquisition tags (D-tags) attached to five individual sperm whales near the Azores, Oliveira et al. (2016) strongly indicated that codas from these

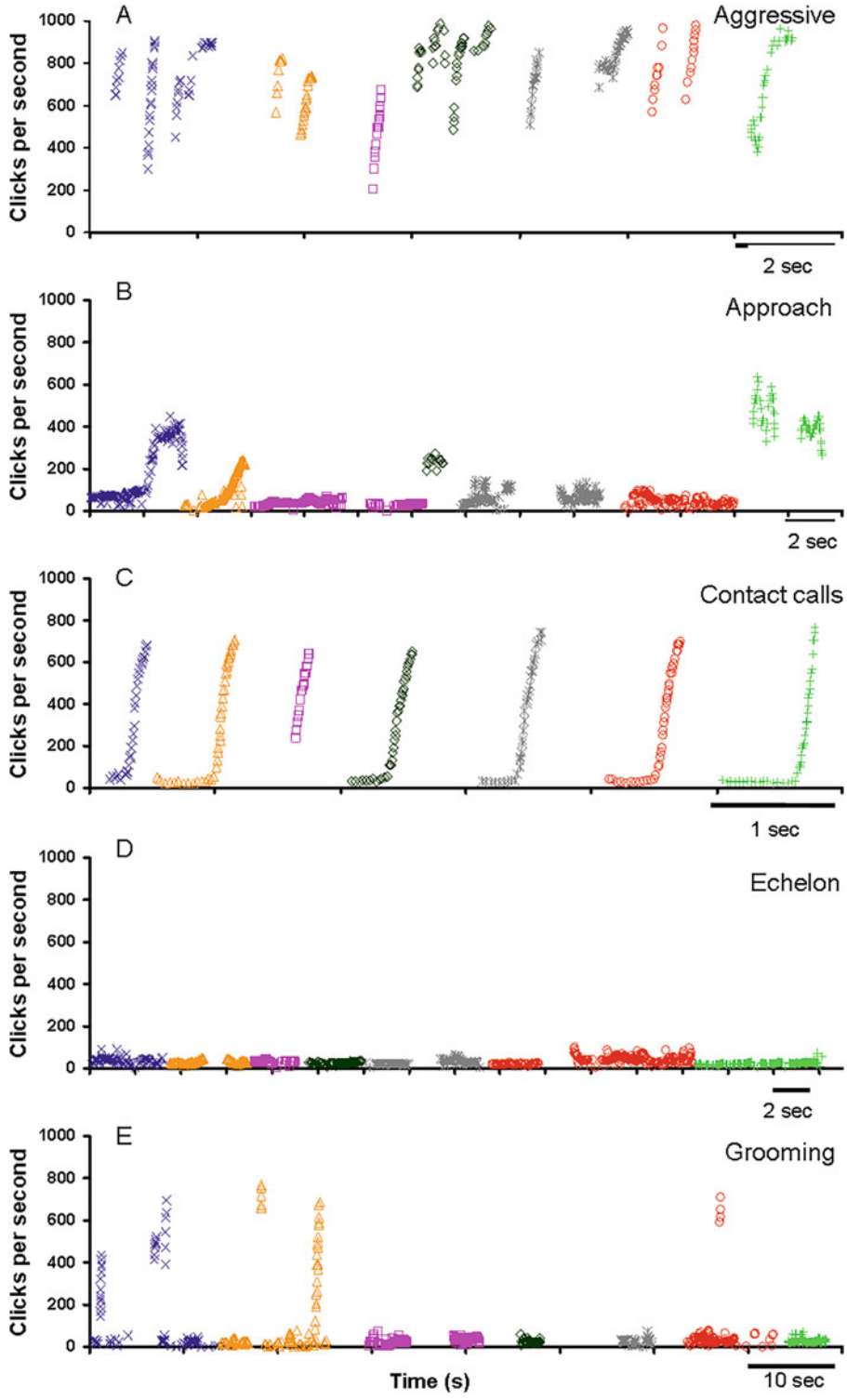



**Fig. 12.23** Use of echolocation click rates by harbor porpoise as communication signals. Five different acoustic behaviors with seven events in each are shown. Note the very rapid increase in click repetition rate up to 1000 clicks/s during aggressive encounters. Reprinted with permission from Taylor & Francis. Clausen KT, Wahlberg M, Beedholm K, Dereuiter S, Madsen PT, Click communication in harbor porpoises (*Phocoena phocoena*). Bioacoustics 20:1–28; https://www.tandfonline.com/doi/abs/10.1080/09524622.2011.9753630. © Taylor & Francis, 2011. All rights reserved

sperm whales contained individual identification information. Some of the patterns can be distinct from one area to another while others, like the five-click coda, occurred in geographically widespread social units. We have yet to reach a detailed understanding of the use of codas by sperm whales, but codas may carry specific behavioral information from individual sperm whales.

Sperm whale coda-clicks resemble biosonar-clicks (Fig. 12.12) and the same basic mechanism likely underlies the production of both. However, whereas the biosonar-click largely bypasses the distal air sac, reducing the strength of back reflections (P1 etc. in Fig. 12.12), the (Po) of the coda-click seems to exit the rostrum more dorsally (see Fig. 12.12). It thus hits a larger portion of the distal air sac and reflects to a larger extent back to the frontal air sac producing the P1. This difference is indicated by the smaller dB difference between the Po and P1 components for coda clicks relative to biosonar clicks (Fig. 12.12). The large muscle and tendon layer between the dorsal edges of the cranium to the tip of the rostrum could play a role in directing the click. The initial coda click (Po) is lower in frequency and intensity than the biosonar click (Fig. 12.12, relative amplitude values). The intervals between repetitions of a coda click match those of a biosonar click from the same animal (Fig. 12.12b) and reflect the distance between the distal (Di) and frontal (Fr) air sacs (see Fig. 12.12). The properties of the coda clicks make them more suited for close-range and less directional communication than the more intense, higher frequency biosonar clicks (Fig. 12.13).

Whether echolocation signals serve a role for intraspecific communication in birds and insectivores has, to our knowledge, not been studied, but Suthers and Hector (1988) hypothesized that individual differences of the syrinx anatomy, specifically the position of the syringeal membranes, would allow oilbirds to distinguish own from conspecific signals by differences in the spectral characteristics of their clicks.

## 12.9 Summary

To date, highly specialized echolocation systems have evolved in many bat species and in toothed whales. Oilbirds and swiftlets also make use of a cruder type of echolocation, independent of obvious auditory specializations, for orientation when their visual abilities become insufficient. A more complete understanding of echolocation by birds awaits future studies. A form of echo-based orientation may be present in shrews and tenrecs, but the exact extent of its function still needs proper documentation.

Most echolocators use ultrasonic signals, either broadband clicks (including most toothed whales, rousette bats, oilbirds and swiftlets) or, as in most bats, tonal echolocation calls of constant frequency, frequency-modulated sweeps, or a combination of these call types. Generally, echolocation signals have high amplitude to promote long-range transmission. Bats and dolphins emit echolocation signals in a narrow beam, a sort of acoustic flashlight, to focus their search. In both bats and dolphins, the repetition rate of signals increases as they approach a target. Bats and dolphins can adjust the frequency and amplitude of their biosonar signals to adapt to noisy ambient conditions. Most echolocators do not broadcast and receive echolocation signals at the same time but separate the outgoing pulse from the echo in time to minimize the masking of faint echoes by the next outgoing signal. However, some families of bats are overlap-tolerant and emit long echolocation signals of constant frequency while listening for Doppler-shifted echoes returned by prey items.

Hearing anatomy, physiology, and abilities in bats and dolphins have been well-studied. Bats have a tragus and grooves in their pinnae that aid in signal reception and directional hearing. In contrast, dolphins do not have pinnae but have evolved asymmetrical skull bones that aid in directional hearing. Some bats emit echolocation signals through their nose and have elaborate nose-leafs while others are open-mouth

echolocators. Bats produce their echolocation sounds in the larynx. Dolphins emit echolocation sounds through the melon within their forehead and from here into the water. They have phonic lips in their nasal passage to produce their echolocation clicks and communication whistles.

A primary advantage of echolocation is allowing animals to operate and orient in situations where light is uncertain, unpredictable, or plain absent. But as with other sensory capacities, echolocation often does not stand alone. The cross-modal sensory interactions between echolocation and sensory abilities such as touch, olfaction, and vision, is an area awaiting further exploration.

Information leakage is a primary disadvantage of echolocation. The signals used in echolocation are audible to many other animals, such as competing conspecifics, predators, and prey. The evolutionary arms race between echolocating bats and some insect prey is a classic example of predator–prey co-evolution. Signals used in echolocation also can function in communication, as shown in echolocating bats and toothed whales.

Both bats and odontocetes are affected by anthropogenic activities, as exemplified by the high mortality experienced by some bat species from wind turbines and incidents of drowning, for example, in porpoises accidentally entangled in stationary gillnets. Anthropogenic sound sources like road or shipping noise may interfere with efficient foraging in bats and toothed whales and seismic explosions used for offshore oil exploration can affect the behavior of toothed whales and other marine mammals. Echolocating birds are also affected by humans, for example, from poaching or nest collecting and habitat-destructive mining activity. Gaining an increased understanding of echolocation behavior in these animals could have important implications for such issues and for wildlife management in general.

## 12.10 Additional Resources

For a more in-depth view of bat echolocation, we strongly recommend Griffin's book *Listening in the Dark*. While now more than 60 years old, the original observations and insights detailed by Griffin (1958) are still very much to the point and relevant today. The Springer Handbook of Auditory Research volumes *Hearing by Bats, Bat Bioacoustics, Hearing by Whales and Dolphins,* and *Biosonar* are also highly recommended as they hold much more detail than the present description. Finally, Thomas, Moss, and Vater edited a book on *Echolocation in Bats and Dolphins* in 2002.

**Acknowledgments** We dedicate this chapter to Dr. Annemarie Surlykke, who made substantial contributions to the field of bioacoustics in insects and in echolocating bats. She was one of the first women scientists to concentrate her research in the area of bioacoustics, which requires a multi-disciplinary understanding of biology, acoustics, physics, animal behavior, and electrical engineering.

We appreciate the careful reviews of sections 5 and 8 by Mats Amundin, Senior Advisor Kolmårdens Djurpark and Guest Prof. Linkoping University, Sweden; Professor Peter T. Madsen, Department of Bioscience, Aarhus University, Denmark; and Associate Professor Magnus Wahlberg, Institute of Biology, University of Southern Denmark, Odense, Denmark. We acknowledge and appreciate the initial outline of this chapter by now deceased Jeanette Thomas.

## References

Amundin M (1991) Sound production in odontocetes with emphasis on the harbour porpoise *Phocoena phocoena*. Stockholm University, Stockholm

Andersen SH, Amundin M (1976) Possible predator-related adaption of sound production and hearing in the harbour porpoise (*Phoconea phocoena*). Aquat Mamm 4(2):56–57

Antunes R, Schulz T, Gero S, Whitehead H, Gordon J, Rendell L (2011) Individual distinctive acoustic features in sperm whale codas. Anim Behav 81(4): 723–730

Arranz P, Aguilar de Soto N, Madsen PT, Brito A, Bordes F, Johnson MP (2011) Following a foraging fish-finder: diel habitat use of Blainville's beaked whales revealed by echolocation. PLoS One 6(12). https://doi.org/10.1371/journal.pone.0028353

Au WWL (1993) The sonar of dolphins. Springer, New York

Au WWL (2015) History of dolphin biosonar research. Acoust Tod 11(4):10–17

Au WWL, Simmons JA (2007) Echolocation in dolphins and bats. Phys Tod 2007:40–45

Au WWL, Suthers RA (2014) Production of biosonar signals: structure and form. In: Surlykke A, Nachtigall PE, Fay RR, Popper AN (eds) Biosonar. Springer, New York, pp 61–105. https://doi.org/10.1007/978-1-4614-9146-0_3

Au WWL, Charder DA, Penner RH, Scronce BL (1985) Demonstration of adaptation in beluga whale echolocation signals. J Acoust Soc Am 77:726–730

Au WWL, Moore PWB, Pawloski D (1986) Echolocation transmitting beam of the Atlantic bottlenose dolphin. J Acoust Soc Am 80:688–691

Au WWL, Kastelein RA, Rippe T, Schooneman NM (1999) Transmission beam pattern and echolocation signals of a harbor porpoise (*Phocoena phocoena*). J Acoust Soc Am 106:3699–3705

Au WWL, Kastelein RA, Benoit-Bird KJ, Cranford TW, McKenna MF (2006) Acoustic radiation from the head of echolocating harbor porpoises (*Phocoena phocoena*). J Exp Biol 209:2726–2733

Baumann-Pickering S, Wiggins SM, Roth EH, Roch MA, Schnitzler HU, Hildebrand JA (2010) Echolocation signals of a beaked whale at Palmyra atoll. J Acoust Soc Am 127(6):3790–3799. https://doi.org/10.1121/1.3409478

Blomkvist C, Amundin M (2004) High-frequency burst-pulse sounds in agonistic/aggressive interactions in bottlenose dolphins, *Tursiops truncatus*. In: Thomas JA, Moss CF, Vater M (eds) Echolocation in bats and dolphins. University of Chicago Press, Chicago, pp 425–431

Boonman AM, Jones G (2002) Intensity control during target approach in echolocating bats; stereotypical sensori-motor behaviour in Daubenton's bats, *Myotis daubentonii*. J Exp Biol 205:2865–2874

Bosque C, Ramirez R, Rodriguez D (1995) The diet of the oilbird in Venezuela. Ornitol Neotrop 6:67–80

Brinkløv S, Fenton MB, Ratcliffe JM (2013) Echolocation in oilbirds and swiftlets. Front Physiol 4:123

Brinkløv S, Elemans CPH, Ratcliffe JM (2017) Oilbirds produce echolocation signals beyond their best hearing range and adjust signal design to natural light conditions. R Soc Open Sci 4(5):17025

Bruns V, Schmieszek E (1980) Cochlear innervation in the greater horseshoe bat - demonstration of an acoustic fovea. Hear Res 3(1):27–43. https://doi.org/10.1016/0378-5955(80)90006-4

Buchler ER (1976) The use of echolocation by the wandering shrew (*Sorex vagrans*). Anim Behav 24:858–873

Burgin CJ, Colella JP, Kahn PL, Upham NS (2018) How many species of mammals are there? J Mammal 99(1):1–14. https://doi.org/10.1093/jmammal/gyx147

Caruso F, Sciacca V, Bellia G, De Domenico E, Larosa G, Papale E et al (2015) Size distribution of sperm whales acoustically identified during long term deep-sea monitoring in the Ionian Sea. PLoS One 10(12):e0144503. https://doi.org/10.1371/journal.pone.0144503

Chantler P, Wells DR, Schuchmann KL (1999) Family Apodidae (swifts). In: Hoyo D, Elliott S (eds) Handbook of the birds of the world, barn owls to hummingbirds, vol 5. Lynx, Barceloa, pp 388–457

Clausen KT, Wahlberg M, Beedholm K, Dereuiter S, Madsen PT (2010) Click communication in harbour porpoises (*Phocoena phocoena*). Bioacoustics 20:1–28

Coles RB, Konishi M, Pettigrew JD (1987) Hearing and echolocation in the Australian Grey swiftlet, *Collocalia spodiopygia*. J Exp Biol 129:365–371

Coles RB, Guppy A, Anderson ME, Schlegel P (1989) Frequency sensitivity and directional hearing in the gleaning bat, *Plecotus auritus* (Linnaeus 1758). J Comp Physiol A 165:269–280

Cranford TW, Amundin M, Norris KS (1996) Functional morphology and homology in the Odontocete nasal complex: implications for sound generation. J Morphol 228:223–285

Cranford TW, McKenna MF, Soldevilla MS, Wiggins SM, Goldbogen JA, Shadwick RE, Krysl P, Leger JA, Hildebrand JA (2008) Anatomic geometry of sound transmission and reception in Cuvier's beaked whale (*Ziphius cavirostris*). Anat Rec 291:353–378

Cranford TW, Elsberry WR, Van Bonn WG, Jeffress JA, Chaplin MS, Blackwood DJ, Carder DA, Kamolnick T, Todd MA, Ridgway SH (2011) Observation and analysis of sonar signal generation in the bottlenose dolphin (*Tursiops truncatus*): evidence for two sonar sources. J Exp Mar Biol Ecol 407(1):81–96

Cranford TW, Trijoulet V, Smith CR, Krysl P (2014) Validation of a vibroacoustic finite element model using bottlenose dolphin simulations: the dolphin biosonar beam is focused in stages. Bioacoustics 23(2):161–194

Cranford TW, Amundin M, Krysl P (2015) Sound production and sound reception in Delphinoids. In: Johnson CM, Herzing DL (eds) Dolphin communication and cognition. Past, present, and future. MIT Press, Boston, MA, pp 19–48

Culik BM (2011) Odontocetes - the toothed whales, CMS Technical Series No. 24, vol 24. United Nations Environmental Program, Bonn, Germany

Dalland JI (1965) Hearing sensitivity in bats. Science 150:1185–1186

Dawson SM (1991) Clicks and communication: the behavioural and social contexts of Hector's dolphin vocalizations. Ethology 88:265–276

Denzinger A, Schnitzler HU (2013) Bat guilds, a concept to classify the highly diverse foraging and echolocation behaviors of microchiropteran bats. Front Physiol 4. https://doi.org/10.3389/fphys.2013.00164

Dooling RJ (1980) Behavior and psychophysics of hearing in birds. In: Popper AN, Fay RR (eds) Comparative studies of hearing in vertebrates. Springer, New York, pp 261–288

Dormer KJ (1979) Mechanism of sound production and air recycling in delphinids: cineradiographic evidence. J Acoust Soc Am 65(1):229–239

Drexl M, Faulstich MH, Von Stebut B, Radtke-Schuller S, Kössl M (2003) Distortion product otoacoustic emissions and auditory evoked potentials in the

hedgehog tenrec, *Echinops telfairi*. J Assoc Res Otolaryngol 4:555–564
Elemans CPH, Mead AF, Jakobsen L, Ratcliffe JM (2011) Superfast muscles set maximum call rate in echolocating bats. Science 333:1885–1888
Fais M, Johnson M, Wilson M, Aguilar Soto N, Madsen PT (2016) Sperm whale predator-prey interactions involve chasing and buzzing, but no acoustic stunning. Sci Rep 6:28562:1–13. https://doi.org/10.1038/srep28562
Falk B, Williams T, Aytekin M, Moss CF (2011) Adaptive behavior for texture discrimination by the free-flying big brown bat, *Eptesicus fuscus*. J Comp Physiol A 197(5):491–503
Fay RR (1988) Hearing in vertebrates: a psychophysics databook. Hill-Fay Associates, Winnetka, IL
Fenton MB (1975) Acuity of echolocation in *Collocalia hirundinacea* (Aves: Apodidae), with comments on the distributions of echolocating swiftlets and molossid bats. Biotropica 7:1–7
Fenton MB (1995) Natural history and biosonar signals. In: Popper AN, Fay RR (eds) Hearing by bats, Springer handbook of auditory research, vol 5. Springer, New York, pp 37–86
Fenton MB, Faure PA, Ratcliffe JM (2012) Evolution of high duty cycle echolocation in bats. J Exp Biol 215(17):2935–2944
Finneran JJ, Branstetter BK, Houser DS, Moore PW, Mulsow J, Martin C, Perisho S (2014) High-resolution measurement of a bottlenose dolphin's (*Tursiops truncatus*) biosonar transmission beam pattern in the horizontal plane. J Acoust Soc Am 136(4):2025–2038
Forsman KA, Malmquist MG (1988) Evidence for echolocation in the common shrew, *Sorex araneus*. J Zool Soc Lond 216:655–662
Fullard JH, Barclay RMR, Thomas DW (1993) Echolocation in free-flying Atiu swiftlets (*Aerodramus sawtelli*). Biotropica 25:334–339
Galatius A, Olsen MT, Steeman ME, Racicot RA, Bradshaw CD, Kyhn L, Miller LA (2019) Raising your voice: evolution of narrow band high frequency signals in odontocetes. Biol J Linn Soc 126:213–224. https://doi.org/10.1093/biolinnean/bly194
Gardiner JD, Dimitriadis G, Sellers WI, Codd JR (2008) The aerodynamics of big ears in the brown long-eared bat *Plecotus auritus*. Acta Chiropterol 10(2):313–321. https://doi.org/10.3161/150811008x414881
Ghose K, Moss CF (2003) The sonar beam pattern of a flying bat as it tracks tethered insects. J Acoust Soc Am 114(2):1120–1131
Goldman LJ, Henson OW (1977) Prey recognition and selection by the constant frequency bat, *Pteronotus p. parnellii*. Behav Ecol Sociobiol 2:411–419
Gonzalez-Terrazas TP, Martel C, Milet-Pinheiro P, Ayasse M, Kalko EKV, Tschapka M (2016) Finding flowers in the dark: nectar-feeding bats integrate olfaction and echolocation while foraging for nectar. R Soc Open Sci 3(8). https://doi.org/10.1098/rsos.160199
Gould E (1965) Evidence for echolocation in the Tenrecidae of Madagascar. Proc Am Philos Soc 109:352–360
Gould E, Negus NC, Novick A (1964) Evidence for echolocation in shrews. J Exp Zool 156:19–37
Griffin DR (1944) Echolocation by blind men, bats and radar. Science 100:589–590
Griffin DR (1953) Acoustic orientation in the oilbird, Steatornis. Proc Natl Acad Sci USA 39:884–893
Griffin DR (1958) Listening in the dark, 2nd edn. Cornell University, New York
Griffin DR, Suthers RA (1970) Sensitivity of echolocation in cave swiftlets. Biol Bull 139:365–371
Griffin DR, Thompson T (1982) Echolocation by cave swiftlets. Behav Ecol Sociobiol 10:119123
Griffin DR, Webster FA, Michael CR (1960) The echolocation of flying insects by bats. Anim Behav 8:141–154
Griffin DR, Friend JH, Webster FA (1965) Target discrimination by the echolocation of bats. J Exp Zool 158:155–168
Grunwald JE, Schornich S, Wiegrebe L (2004) Classification of natural textures in echolocation. Proc Natl Acad Sci USA 101(15):5670–5674. https://doi.org/10.1073/pnas.0308029101
Guppy A, Coles RB (1988) Acoustical and neural aspects of hearing in the Australian gleaning bats, *Macroderma gigas* and *Nyctophilus gouldi*. J Comp Physiol A 162(5):653–668. https://doi.org/10.1007/Bf01342641
Hartley DJ (1992a) Stabilization of perceived echo amplitudes in echolocating bats. I. Echo detection and automatic gain control in the big brown bat, *Eptesicus fuscus*, and the fishing bat, *Noctilio leporinus*. J Acoust Soc Am 91:1120–1132
Hartley DJ (1992b) Stabilization of perceived echo amplitudes in echolocating bats. II. The acoustic behavior of the big brown bat, *Eptesicus fuscus*, when tracking moving prey. J Acoust Soc Am 91:1133–1149
Hartley DJ, Suthers RA (1987) The sound emission pattern and the acoustical role of the noseleaf in the echolocating bat, *Carollia perspicillata*. J Acoust Soc Am 82:1892–1900
Hartley DJ, Suthers RA (1989) The sound emission pattern of the echolocating bat, *Eptesicus fuscus*. J Acoust Soc Am 85:1348–1351
He K, Liu Q, Xu D-M, Qi F-Y, Bai J, He S-W, Chen P, Zhou X, Cai W-Z, Chen Z-L, Jiang X-L, Shi P (2021) Echolocation in soft-furred tree mice. Science 372:1–10. https://doi.org/10.1126/science.aay1513
Henson OW Jr (1965) The activity and function of the middle-ear muscles in echo-locating bats. J Physiol 180(4):871–887. https://doi.org/10.1113/jphysiol.1965.sp007737
Herzing DL (2000) Acoustics and social behavior of wild dolphins: implications for a sound society. In: Au WWL, Popper AN, Fay RR (eds) Hearing by whales

and dolphins, Hearing by whales and dolphins, vol 12. Springer, New York, pp 225–272

Hiryu S, Mora EC, Riquimaroux H (2016) Behavioral and physiological bases for doppler shift compensation by echolocating bats. In: Fenton M, Grinnell A, Popper A, Fay R (eds) Bat Bioacoustics, Springer handbook of auditory research, vol 54. Springer, New York, NY. https://doi.org/10.1007/978-1-4939-3527-7_9

Holderied MW, Korine C, Fenton MB, Parsons S, Robson S, Jones G (2005) Echolocation call intensity in the aerial hawking bat *Eptesicus bottae* (Vespertilionidae) studied using stereo videogrammetry. J Exp Biol 208:1321–1327

Houston RD, Boonman AM, Jones G (2004) Do echolocation signal parameters restrict bats' choice of prey? In: Thomas JA, Moss CF, Vater M (eds) Echolocation in bats and dolphins. Chicago University Press, Chicago, pp 339–345

Huggenberger S, Rauschmann MA, Vogl TJ, Oelschläger HHA (2009) Functional morphology of the nasal complex in the harbor porpoise (*Phocoena phocoena* L.). Anat Rec 292:902–920

Hulgard K, Moss CF, Jakobsen L, Surlykke A (2016) Big brown bats (*Eptesicus fuscus*) emit intense search calls and fly in stereotyped flight paths as they forage in the wild. J Exp Biol 219(3):334–340. https://doi.org/10.1242/jeb.128983

Iwaniuk AN, Clayton DH, Wylie DR (2006) Echolocation, vocal learning, auditory localization and the relative size of the avian auditory midbrain nucleus (MLd). Behav Brain Res 167:307–317

Jakobsen L, Surlykke A (2010) Vespertilionid bats control the width of their biosonar sound beam dynamically during prey pursuit. PNAS 107(31):13930–13935

Jakobsen L, Brinklov S, Surlykke A (2013a) Intensity and directionality of bat echolocation signals. Front Physiol 4:89. https://doi.org/10.3389/fphys.2013.00089

Jakobsen L, Ratcliffe JM, Surlykke A (2013b) Convergent acoustic field of view in echolocating bats. Nature 493(7430):93–96. https://doi.org/10.1038/nature11664

Jakobsen L, Olsen MN, Surlykke A (2015) Dynamics of the echolocation beam during prey pursuit in aerial hawking bats. Proc Natl Acad Sci USA 112(26): 8118–8123. https://doi.org/10.1073/pnas.1419943112

Jen PHS, Suga N (1976) Coordinated activities of middle-ear and laryngeal muscles in echolocating bats. Science 191:950–952

Johansson LC, Hakansson J, Jakobsen L, Hedenstrom A (2016) Ear-body lift and a novel thrust generating mechanism revealed by the complex wake of brown long-eared bats (*Plecotus auritus*). Sci Rep 6:24886. https://doi.org/10.1038/srep24886

Johnson M, Madsen PT, Zimmer WMX, Aguilar de Soto N, Tyack PL (2006) Foraging Blainville's beaked whales (*Mesoplodon densirostris*) produce distinct click types matched to different phases of echolocation. J Exp Biol 209:5038–5050

Kalko EKV, Schnitzler H-U (1989) The echolocation and hunting behavior of Daubenton's bat, *Myotis daubentoni*. Behav Ecol Sociobiol 24:225–238

Kalko EKV, Schnitzler H-U (1993) Plasticity in echolocation signals of European pipistrelle bats in search flight: implications for habitat use and prey detection. Behav Ecol Sociobiol 33:415–428

Kastelein RA, Hoek L, de Jong CAF, Wensveen PJ (2010) The effect of signal duration on the underwater detection thresholds of a harbor porpoise (*Phocoena phocoena*) for single frequency-modulated tonal signals between 0.25 and 160 kHz. J Acoust Soc Am 128(5): 3211–3222. https://doi.org/10.1121/1.3493435_

Kloepper LN, Nachtigall PE, Donahue MJ, Breese M (2012) Active echolocation beam focusing in the false killer whale, *Pseudorca crassidens*. J Exp Biol 215:1306–1312. https://doi.org/10.1242/jeb.066605

Koay G, Heffner RS, Heffner HE (1998) Hearing in a Megachiropteran fruit bat (*Rousettus aegyptiacus*). J Comp Psychol 112(4):371–382

Koblitz JC, Wahlberg M, Stilz P, Madsen PT, Beedholm K, Schnitzler HU (2012) Asymmetry and dynamics of a narrow sonar beam in an echolocating harbor porpoise. J Acoust Soc Am 131(3):2315–2324. https://doi.org/10.1121/1.3683254

Konishi M, Knudsen EI (1979) The oilbird: hearing and echolocation. Science 204:425–427

Kössl M, Vater M (1995) Cochlear structure and function in bats. In: Popper AN, Fay RR (eds) Hearing by bats, vol 5. Springer, New York, pp 191–234

Kounitsky P, Rydell J, Amichai E, Boonman A, Eitan O, Weiss AJ, Yovel Y (2015) Bats adjust their mouth gape to zoom their biosonar field of view. Proc Natl Acad Sci USA 112(21):6724–6729. https://doi.org/10.1073/pnas.1422843112

Kubke MF, Massoglia DP, Carr CE (2004) Bigger brains or bigger nuclei? Regulating the size of auditory structures in birds. Brain Behav Evol 63:169–180

Kuroda M, Sasaki M, Yamada K, Miki N, Matsuishi T (2015) Tissue physical property of the harbor porpoise *Phocoena phocoena* for investigation of the sound emission process. J Acoust Soc Am 138(3):1451–1456

Kürten L, Schmidt U (1982) Thermoperception in the common vampire bat (*Desmodus rotundus*). J Comp Physiol 146:223–228. https://doi.org/10.1007/BF00610241

Ladegaard M, Jensen FH, Beedholm K, da Silva VMF, Madsen PT (2017) Amazon river dolphins (*Inia geoffrensis*) modify biosonar output level and directivity during prey interception in the wild. J Exp Biol 220: 2654–2665. https://doi.org/10.1242/jeb.159913

Lammers MO, Au WWL, Aubauer R, Nachtigall PE (2004) A comparative analysis of the pulsed emissions of free-ranging Hawaiian spinner dolphins (*Stenella longirostris*). In: Thomas JA, Moss CF, Vater M (eds) Echolocation in bats and dolphins. University of Chicago Press, Chicago, pp 414–419

Lawrence BD, Simmons JA (1982a) Echolocation in bats: the external ear and perception of the vertical position of targets. Science 218:481–483

Lawrence BD, Simmons JA (1982b) Measurements of atmospheric attenuation at ultrasonic frequencies and the significance for echolocation by bats. J Acoust Soc Am 71(3):585–590. https://doi.org/10.1121/1.387529

Lazure L, Fenton MB (2011) High duty-cycle echolocation and prey detection by bats. J Exp Biol 214:1131–1137

Lewanzik D, Goerlitz HR (2018) Continued source level reduction during attack in the low-amplitude bat *Barbastella barbastellus* prevents moth evasive flight. Funct Ecol 32(5):1251–1261. https://doi.org/10.1111/1365-2435.13073

Li S, Nachtigall PE, Breese M (2011) Dolphin hearing during echolocation: evoked potential responses in an Atlantic bottlenose dolphin (*Tursiops truncatus*). J Exp Biol 214:2027–2035. https://doi.org/10.1242/jeb.053397

Linnenschmidt M, Beedholm K, Wahlberg M, Kristensen JH, Nachtigall PE (2012) Keeping returns optimal: gain control exerted through sensitivity adjustments in the harbour porpoise auditory system. Proc Royal Soc B 279:2237–2465

Linnenschmidt M, Teilmann J, Akamatsu T, Dietz R, Miller LA (2013) Biosonar, dive and foraging activity of satellite tracked harbour porpoises (*Phocoena phocoena*). Mar Mamm Sci 29(2):E77–E97

Liu Z, Qi F-Y, Zhou X, Ren H-Q, Shi P (2014) Parallel sites implicate functional convergence of the hearing gene *Prestin* among echolocating mammals. Mol Biol Evol 31(9):2415–2424. https://doi.org/10.1093/molbev/msu194

Long GR, Schnitzler HU (1975) Behavioral audiograms for the bat, *Rhinolophus ferrumequinum*. J Comp Physiol A 100:211–220

Madsen PT, Surlykke A (2014) Echolocation in air and water. In: Surlykke A, Nachtigall PE, Fay RR, Popper AN (eds) Biosonar, Springer handbook of auditory research, vol 51. Springer, New York, pp 257–304. https://doi.org/10.1007/978-1-4614-9146-0

Madsen PT, Carder DA, Beedholm K, Ridgway SH (2005) Porpoise clicks from a sperm whale nose – convergent evolution of 130 kHz pulses in toothed whale sonars? Bioacoustics 15:195–206

Madsen PT, Wilson M, Johnson M, Hanlon RT, Bocconcelli A, Aguilar de Soto N, Tyack PL (2007) Clicking for calamari: toothed whales can echolocate squid *Loligo pealeii*. Aquat Biol 1:141–150

Madsen PT, Wisniewska D, Beedholm K (2010) Single source sound production and dynamic beam formation in echolocating harbour porpoises (*Phocoena phocoena*). J Exp Biol 213:3105–3110. https://doi.org/10.1242/jeb.044420

Madsen PT, Lammers MO, Wisniewska D, Beedholm K (2013) Nasal sound production in echolocating delphinids (*Tursiops truncatus* and *Pseudorca crassidens*) is dynamic, but unilateral: clicking on the right side and whistling on the left side. J Exp Biol 216(21):4091–4102. https://doi.org/10.1242/jeb.091306

Manley GA (1990) Peripheral hearing mechanisms in reptiles and birds. Springer, Berlin

Mann DA, Lu Z, Popper AN (1997) Ultrasound detection by a teleost fish. Nature 389:341–341

Martin GR, Rojas LM, Ramírez Y, McNeil R (2004) The eyes of oilbirds (*Steatornis caripensis*): pushing the limits of sensitivity. Naturwissenschaften 91:26–29

Masters WM, Moffat AJM, Simmons JA (1985) Sonar tracking of horizontally moving targets by the big brown bat *Eptesicus fuscus*. Science 228:1331–1333

Matsuta N, Hiryu S, Fujioka E, Yamada Y, Riquimaroux H, Watanabe Y (2013) Adaptive beam-width control of echolocation sounds by CF-FM bats, *Rhinolophus ferrumequinum nippon*, during prey-capture flight. J Exp Biol 216(Pt 7):1210–1218. https://doi.org/10.1242/jeb.081398

Miller LA (1983) IV.3 How insects detect and avoid bats. In: Huber F, Markl H (eds) Neuroethology and behavioral physiology. Springer, Berlin, pp 251–266

Miller LA (2010) Prey capture by harbor porpoises (*Phocoena phocoena*): a comparison between echolocators in the field and in captivity. J Mar Acoust Soc Jpn 37(3):156–168

Miller LA, Surlykke A (2001) How some insects detect and avoid being eaten by bats: the tactics and counter tactics of prey and predator. Bioscience 51:570–581

Miller LA, Wahlberg M (2013) Echolocation by the harbour porpoise: life in coastal waters. Front Integr Physiol 4(52):1–6. https://doi.org/10.3389/fphys.2013.00052

Møhl B, Wahlberg M, Madsen PT, Heerfordt A, Lund A (2003) The monopulsed nature of sperm whale clicks. J Acoust Soc Am 114(2):1143–1154

Moore P, Popper AN (2019) Heptuna's contributions to biosonar. Acoust Tod 15(1):44–52

Moore PWB, Dankiewicz LA, Houser DS (2008) Beamwidth control and angular target detection in an echolocating bottlenose dolphin (*Tursiops truncatus*). J Acoust Soc Am 124:3324–3332. https://doi.org/10.1121/1.2980453

Morisaka T, Connor RC (2007) Predation by killer whales (*Orcinus orca*) and the evolution of whistle loss and narrow-band high frequency clicks in odontocetes. Evol Biol 20. https://doi.org/10.1111/j.1420-9101.2007.01336.x

Motoi K, Sumiya M, Fujioka E, Hiryu S (2017) Three-dimensional sonar beam-width expansion by Japanese house bats (*Pipistrellus abramus*) during natural foraging. J Acoust Soc Am 141(5):EL439. https://doi.org/10.1121/1.4981934

Nachtigall PE, Supin AY (2008) A false killer whale adjusts its hearing when it echolocates. J Exp Biol 211:1714–1718

Nachtigall PE, Yuen MML, Mooney TA, Taylor KA (2005) Hearing measurements from a stranded infant Risso's dolphin, *Grampus griseus*. J Exp Biol 208:4181–4188

Nachtigall PE, Mooney TA, Taylor KA, Miller LA, Rasmussen MH, Akamatsu T, Teilmann J, Linnenschmidt M, Vikingsson GA (2008) Shipboard measurements of the hearing of the white-beaked dolphin, *Lagenorhynchus albirostris*. J Exp Biol 211:642–647

Neuweiler G (2000) The biology of bats (trans: Covey E). Oxford University Press, Oxford

Noren DP, Holt MM, Dunkin RC, Williams TM (2017) Echolocation is cheap for some mammals: dolphins conserve oxygen while producing high-intensity clicks. J Exp Mar Biol Ecol 495:103–109

Norris KS, Møhl B (1983) Can odontocetes debilitate prey with sound? Am Nat 122(1):85–104
Norris KS, Prescott JH, Asa-Dorian PV, Perkins P (1961) An experimental demonstration of echo-location behavior in the porpoise, *Tursiops truncatus* (Montagu). Biol Bull 120:163–176
Norris KS, Dormer KJ, Pegg J, Liese GJ (1971) The mechanism of sound production and air recycling in porpoises: a preliminary report. In: Paper presented at the Proceedings of the eighth conference on the biological sonar of diving mammals. Menlo Park, California
Nørum U, Brinkløv S, Surlykke A (2012) New model for gain control of signal intensity to object distance in echolocating bats. J Exp Biol 215(17):3045–3054
Novick A (1959) Acoustic orientation in the cave swiftlet. Biol Bull 117:497–503
Obrist MK, Fenton MB, Eger JL, Schlegel PA (1993) What ears do for bats: a comparative study of pinna sound pressure transformation in Chiroptera. J Exp Biol 180:119–152
Oliveira C, Wahlberg M, Silva MA, Johnson M, Antunes R, Wisniewska D, Fais A, Madsen PT (2016) Sperm whale codas may encode individuality as well as clan identity. J Acoust Soc Am 139(5): 2860–2869
Pedersen SC (1993) Cephalometric correlates of echolocation in the chiroptera. J Morphol 218(1):85–98. https://doi.org/10.1002/jmor.1052180107
Pollak GD (1988) Time is traded for intensity in the bats auditory-system. Hear Res 36(2–3):107–124. https://doi.org/10.1016/0378-5955(88)90054-8
Popov AV, Supin AY, Wang D, Wang K (2006) Nonconstant quality of auditory filters in the porpoises, *Phocoena phocoena* and *Neophocaena phocaenoides* (Cetacea, Phocoenidae). J Acoust Soc Am 119(5): 3173–3180
Price JJ, Johnson KP, Clayton DH (2004) The evolution of echolocation in swiftlets. J Avian Biol 35:135–143
Pye JD (1980) Echolocation signals and echoes in air. In: Busnel R-G, Fish JF (eds) Animal sonar systems. Plenum Press, New York, pp 309–353
Pye JD (1993) Is fidelity futile? The 'true' signal is illusory, especially with ultrasound. Bioacoustics 4(4): 271–286
Rasmussen MH, Miller LA (2002) Whistles and clicks from white-beaked dolphins, *Lagenorhynchus albirostris*, recorded in Faxafloi Bay, Iceland. Aquat Mamm 28:78–89
Rasmussen MH, Atem ACG, Miller LA (2016) Behavioral responses by Icelandic white-beaked dolphins (*Lagenorhynchus albirostris*) to playback sounds. Aquat Mamm 42(3):317–329. https://doi.org/10.1578/AM.42.3.2016.317
Ratcliffe JM, Elemans CPH, Jakobsen L, Surlykke A (2013) How the bat got its buzz. Biol Lett 9(2). https://doi.org/10.1098/rsbl.2012.1031
Ridgway SH, Carter DA (1988) Nasal pressure and sound production in an echolocating white whale, *Delphinaperus leucas*. In: Nachtigall PE, Moore PWB (eds) Animal sonar: processes and performance, NATO ASI Series, vol 156. Plenum Press, New York, pp 53–60
Rossbach KA, Herzing DL (1997) Underwater observations of benthic-feeding bottlenose dolphins (*Tursiops truncatus*) near Grand Bahama Island, Bahamas. Mar Mamm Sci 13(3):498–504
Rydell J, Miller LA, Jensen ME (1999) Echolocation constraints of Daubenton's bat foraging over water. Funct Ecol 13:247–255
Sales GD, Pye JD (1974) Ultrasonic communication by animals. Chapman & Hall, London
Sanchez L, Ohdachi SD, Kawahara A, Echenique-Diaz LM, Maruyama S, Kawata M (2019) Acoustic emissions of *Sorex unguiculatus* (Mammalia: Soricidae): assessing the echo-based orientation hypothesis. Ecol Evol 9(5):2629–2639. https://doi.org/10.1002/ece3.4930
Schmidt S (1988) Evidence for a spectral basis of texture perception in bat sonar. Nature 331:617–619
Schmidt S, Türke B, Vogler B (1983) Behavioural audiogram from the bat *Megaderma lyra* (Geoffroy, 1810; Microchiroptera). Myotis 21–22:62–66
Schnitzler HU (1968) Die Ultraschallortungslaute der Hufeisen-Fledermäuse (Chiroptera-Rhinolophidae) in verschiedenen Orientierungssituationen. Z Vgl Physiol 57:376–408
Schnitzler HU (1973) Control of Doppler-shift compensation in greater horseshoe bat, *Rhinolophus ferrumequinum*. J Comp Physiol 82(1):79–92. https://doi.org/10.1007/Bf00714171
Schnitzler HU, Flieger E (1983) Detection of oscillating target movements by echolocation in the greater horseshoe bat. J Comp Physiol A 153:385–391
Schuller G (1977) Echo delay and overlap with emitted orientation sounds and Doppler-shift compensation in bat, *Rhinolophus ferrumequinum*. J Comp Physiol 114(1):103–114
Schuller G, Pollack GD (1979) Disproportionate frequency representation in the inferior colliculus of horseshoe bats: evidence for an "acoustic fovea". J Comp Physiol 132:47–54
Schwartz C, Tressler J, Keller H, Vanzant M, Ezell S, Smotherman M (2007) The tiny difference between foraging and communication buzzes uttered by the Mexican free-tailed bat, *Tadarida brasiliensis*. J Comp Physiol A 193:853–863
Siemers BM, Schauermann G, Turi H, Von Merten S (2009) Why do shrews Twitter? Communication or simple echo-based orientation. Biol Lett 5:593–596
Simmons JA, Howell DJ, Suga N (1975) Information content of bat sonar echoes. Am Sci 63:204–215
Simmons JA, Fenton MB, O'Farrell MJ (1979) Echolocation and pursuit of prey by bats. Science 203:16–21
Simmons JA, Kick SA, Lawrence BD, Hale C, Bard C, Escudié B (1983) Acuity of horizontal angle discrimination by the echolocating bat, *Eptesicus fuscus*. J Comp Physiol A 153:321–330
Simmons JA, Moss CF, Ferragamo M (1990) Convergence of temporal and spectral information into acoustic images of complex sonar targets perceived by the

echolocating bat, *Eptesicus fuscus*. J Comp Physiol A 166:449–470
Smyth DM, Roberts JR (1983) The sensitivity of echolocation by the Grey Swiftlet *Aerodramus spodiopygius*. Ibis 125:339–345
Snow DW (1961) The natural history of the oilbird, *Steatornis caripensis*, in Trinidad, W. I. Part 1. General behavior and breeding habits. Zoologica 46:27–48
Snow DW (1962) The natural history of the oilbird in Trinidad. W. I. Part II. Population breeding ecology, and food. Zoologica 27:199–221
Song Z, Xu X, Dong J, Xing L, Zhang M, Liu X, Zhang Y, Li S, Berggren P (2015) Acoustic property reconstruction of a pygmy sperm whale (*Kogia breviceps*) forehead based on computed tomography imaging. J Acoust Soc Am 138(5):3129–3137
Sørensen PM, Wisniewska DM, Jensen FH, Johnson M, Teilmann J, Madsen PT (2018) Click communication in wild harbour porpoises (*Phocoena phocoena*). Sci Rep 8:9702. https://doi.org/10.1038/s41598-018-28022-8
Speakman JR, Racey PA (1991) No cost of echolocation for bats in flight. Nature 350:421–423
Starkhammar J, Moore PW, Talmadge L, Houser DS (2011) Frequency-dependent variation in the 2-dimensional beam pattern of an echolocating dolphin. Biol Lett 7:836–839
Stilz WP, Schnitzler HU (2012) Estimation of the acoustic range of bat echolocation for extended targets. J Acoust Soc Am 132:1765–1775. https://doi.org/10.1121/1.4733537
Strother GK, Mogus M (1970) Acoustical beam patterns for bats: some theoretical considerations. J Acoust Soc Am 48(6):1430–1432
Suga N, Jen PH-S (1975) Peripheral control of acoustic signals in the auditory system of echolocating bats. J Exp Biol 62:277–311
Surlykke A, Kalko EKV (2008) Echolocating bats cry out loud to detect their prey. PLoS One 3(4):e2036
Surlykke A, Pedersen SB, Jakobsen L (2009a) Echolocating bats emit a highly directional sonar sound beam in the field. Proc R Soc B 276:853–860
Surlykke A, Ghose K, Moss CF (2009b) Acoustic scanning of natural scenes by echolocation in the big brown bat, *Eptesicus fuscus*. JEB 212:1011–1020
Suthers RA, Hector DH (1982) Mechanism for the production of echolocating clicks by the Grey Swiftlet, *Collocalia spodiopygia*. J Comp Physiol A 148:457–470
Suthers RA, Hector DH (1985) The physiology of vocalization by the echolocating oilbird, *Steatornis caripensis*. J Comp Physiol A 156:243–266
Suthers RA, Hector DH (1988) Individual variation in vocal tract resonance may assist oilbirds in recognizing echoes of their own sonar clicks. In: Nachtigall PE, Moore PWB (eds) Animal sonar: processes and performance. Plenum Press, New York, pp 87–91
Szymanski MD, Bain DE, Kiehi K, Pennington S, Wong S, Henry KR (1999) Killer whale (*Orcinus orca*) hearing: auditory brainstem response and behavioral audiograms. J Acoust Soc Am 106:1134–1141
Teeling E (2009) A molecular and morphological perspective on the evolution of echolocation in bats. J Vertebr Paleontol 29:190a–190a
Teeling EC, Springer MS, Madsen O, Bates P, O'Brien SJ, Murphy WJ (2005) A molecular phylogeny for bats illuminates biogeography and the fossil record. Science 307:580–584
Thiagavel J, Cechetto C, Santana SE, Jakobsen L, Warrant EJ, Ratcliffe JM (2018) Auditory opportunity and visual constraint enabled the evolution of echolocation in bats. Nat Commun 9(1):98. https://doi.org/10.1038/s41467-017-02532-x
Thomas JA, Turl CW (1990) Echolocation characteristics and range detection threshold of a false killer whale (*Pseudorca crassidens*). In: Thomas JA, Kastelein RA (eds) Sensory abilities of cetaceans. Plenum Press, New York, pp 321–334
Thomas JA, Moss CF, Vater M (2004) Echolocation in bats and dolphins. University of Chicago Press, Chicago
Thomassen HA (2005) Swift as sound – design and evolution of the echolocation system in Swiftlets (*Apodidae: Collocaliini*). Leiden University, Leiden
Thomassen HA, Povel GDE (2006) Comparative and phylogenetic analysis of the echo clicks and social vocalisations of swifts and swiftlets (Aves: Apodidae). Biol J Linn Soc 88:631–643
Thomassen HA, Djasim UM, Povel GDE (2004) Echo click design in swiftlets: single as well as double clicks. Ibis 146:173–174
Tomasi TE (1979) Echolocation by the short-tailed shrew, *Blarina brevicauda*. J Mammal 60:751–759
Tomassen HA, Gea S, Maas S, Dirckx JJJ, Decraemer WF, Povel GDE (2007) Do Swiftlets have an ear for echolocation? The functional morphology of Swiftlets' middle ears. Hear Res 225:25–37. https://doi.org/10.1016/j.heares.2006.11.013
Urick RJ (1983) Principles of underwater sound, 3rd edn. McGraw-Hill, New York
Vanderelst D, Peremans H, Razak NA, Verstraelen E, Dimitriadis G (2015) The aerodynamic cost of head morphology in bats: maybe not as bad as it seems. PLoS One 10(5):e0118545. https://doi.org/10.1371/journal.pone.0126061
Verfuss UK, Miller LA, Pilz PKD, Schnitzler HU (2009) Echolocation by two foraging harbour porpoises (*Phocoena phocoena*). J Exp Biol 212:823–834
Wang Z, Zhu T, Xue H, Fang N, Zhang J, Zhang L, Pang J, Teeling EC, Zhang S (2017) Prenatal development supports a single origin of laryngeal echolocation in bats. Nat Ecol Evol 1:0021. https://doi.org/10.1038/s41559-016-0021
Watkins WA, Schevill WE (1977) Sperm whale codas. J Acoust Soc Am 62:1485–1490
Wei C, Au WWL, Ketten DR, Song Z, Zhang Y (2017) Biosonar signal propagation in the harbor porpoise's (*Phocoena phocoena*) head: the role of various

structures in the formation of the vertical beam. J Acoust Soc Am 141(6):4179–4187

Wei C, Au WWL, Ketten DR, Zhang Y (2018) Finite element simulation of broadband biosonar signal propagation in the near- and far-field of an echolocating Atlantic bottlenose dolphin (*Tursiops truncatus*). J Acoust Soc Am 143(5):2611–2620. https://doi.org/10.1121/1.5034464

Weissenbacher P, Wiegrebe L (2003) Classification of virtual objects in the echolocating bat, *Megaderma lyra*. Behav Neurosci 117(4):833–839. https://doi.org/10.1037/0735-7044.117.4.833

Wilson M, Acolas ML, Bégout ML, Madsen PT, Wahlberg M (2008) Allis shad (*Alosa alosa*) exhibit an intensity-graded behavioral response when exposed to ultrasound. J Acoust Soc Am 243:1–5

Wilson M, Wahlberg M, Surlykke A, Madsen PT (2013) Ultrasonic predator-prey interaction in water-convergent evolution with insects and bats in air? Front Physiol 4(June):1–12

Wisniewska DM, Johnson M, Beedholm K, Wahlberg M, Madsen PT (2012) Acoustic gaze adjustments during active target selection in echolocating porpoises. J Exp Biol 215:4358–4373

Wisniewska DM, Ratcliffe JM, Beedholm K, Christensen CB, Johnson M, Koblitz JC, Wahlberg M, Madsen PT (2015) Range-dependent flexibility in the acoustic field of view of echolocating porpoises (*Phocoena phocoena*). eLife. https://doi.org/10.7554/eLife.05651

Wong A, Gall MD (2015) Frequency sensitivity in the auditory periphery of male and female black-capped chickadees (*Poecile atricapillus*). Zoology 118:357–363

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



# The Effects of Noise on Animals

# 13

Christine Erbe, Micheal L. Dent, William L. Gannon, Robert D. McCauley, Heinrich Römer, Brandon L. Southall, Amanda L. Stansbury, Angela S. Stoeger, and Jeanette A. Thomas

## 13.1 Introduction

Noise is ubiquitous in all animal habitats, often at substantial levels (Brumm and Slabbekoorn 2005). Habitats typically contain a myriad of geophysical, biological, and anthropogenic sounds, which constitute the local soundscape (see Chap. 7). Some of these sounds can interfere with the life functions of animals and hence are often referred to as "noise" (American National Standards Institute 2013).

Communication plays a critical role in animals' life functions as it is the foundation for social relationships among animals. However, acoustic communication often is constrained by background noise, which reduces the signal-to-noise ratio (SNR) and thus the signal detection and discrimination success of receivers. In terrestrial habitats, natural, abiotic noise is caused by wind, precipitation, thunder, running water, and seismicity. Birds, frogs, insects, and mammals create biotic noise. In aquatic environments, natural, abiotic noise is caused by wind, precipitation, breaking waves, polar ice break-up, and natural seismic activity. Biotic noise sources include shrimps, fishes, and marine mammals.

Such natural noise has been shown to interfere with sound usage by animals. For example, wind noise might interfere with marine mammal communication, and as a counteraction, humpback whales (*Megaptera novaeangliae*) increase the sound pressure level of their sounds as a function of increasing wind noise level (Dunlop et al. 2014). Also, animals of the same or different species can interfere with sound usage. Snapping shrimp are known to mask toothed whale biosonar (Au et al. 1974, 1985) and harp seals (*Pagophilus groenlandicus*) have been shown to

Jeanette A. Thomas (deceased) contributed to this chapter while at the Department of Biological Sciences, Western Illinois University-Quad Cities, Moline, IL, USA

C. Erbe (✉) · R. D. McCauley
Centre for Marine Science & Technology, Curtin University, Perth, WA, Australia
e-mail: c.erbe@curtin.edu.au; r.mccauley@cmst.curtin.edu.au

M. L. Dent
Department of Psychology, University at Buffalo, SUNY, Buffalo, NY, USA
e-mail: mdent@buffalo.edu

W. L. Gannon
Department of Biology and Graduate Studies, Museum of Southwestern Biology, University of New Mexico, Albuquerque, NM, USA
e-mail: wgannon@unm.edu

H. Römer
Department of Biology, Graz University, Graz, Austria
e-mail: heinrich.roemer@uni-graz.at

B. L. Southall
El Paso Zoo, El Paso, TX, USA
e-mail: brandon.southall@sea-inc.net

A. L. Stansbury
Mammal Communication Laboratory, University of Vienna, Vienna, Austria

A. S. Stoeger
Southall Environmental Associates, Inc., Aptos, CA, USA
e-mail: angela.stoeger-horwath@univie.ac.at

© The Author(s) 2022

C. Erbe, J. A. Thomas (eds.), *Exploring Animal Behavior Through Sound: Volume 1*,
https://doi.org/10.1007/978-3-030-97540-1_13

increase their call repetition to be heard above the chorus of their conspecifics (Serrano and Terhune 2001). Similarly, king penguins (*Aptenodytes patagonicus*; Aubin and Jouventin 1998), zebra finches (*Taeniopygia guttata*; Narayan et al. 2007), and big brown bats (*Eptesicus fuscus*; Warnecke et al. 2015) communicate in a cacophony of conspecific calls. Animals have evolved sound production and reception capabilities in natural biotic and abiotic background noise. However, anthropogenic noise is fairly recent on evolutionary time scales. Researchers have tried to assess whether existing adaptations are sufficient for animals to deal with anthropogenic noise.

Anthropogenic noise in terrestrial environments originates from road traffic, trains, aircraft, industrial sites, energy plants, construction machinery, etc. Anthropogenic noise in aquatic environments originates from recreational boating, commercial shipping, commercial fishing, offshore hydrocarbon and mineral exploration, hydrocarbon production, mineral mining, marine construction, offshore renewable energy production, military activities, etc. Such anthropogenic sounds, in air or water, have distinct "sound signatures," and their contributions to the marine and terrestrial soundscapes are discussed in Chap. 7.

The effects of anthropogenic noise have been studied extensively in humans (Kryter 1994); however, less is known about how human-generated noise affects other animals. Four edited books (Brumm 2013; Popper and Hawkins 2012, 2016; Slabbekoorn et al. 2018a) and some journal special issues (Erbe et al. 2016b, 2019c; Le Prell et al. 2019; Thomsen et al. 2020) compile many examples outlining the effects of noise. The effects of anthropogenic noise on animals are a growing concern, having resulted in an exponential increase in the number of research publications on this topic (Williams et al. 2015).

What are the effects of anthropogenic noise? They can vary from mere auditory sensation, mild and temporary annoyance, brief behavioral changes, temporary avoidance of an area, and masking to long-term changes in the usage of important feeding or breeding areas, prolonged stress, hearing loss, barotrauma (in aquatic species), injury, and ultimately death (Kight and Swaddle 2011). In addition to such direct effects of noise, there may be indirect effects (e.g., when a prey species is impacted, leading to reduced prey availability). The effects of noise do not always have to be negative from the animals' point of view. In some cases, animals actually use anthropogenic sounds to their advantage. For example, the sound of a dumpster lid closing in a campground might indicate a food source to some birds and mammals. Underwater sounds from ships can increase the settlement, growth rate, and absolute growth of biofouling organisms such as bryozoans, oysters, calcareous tubeworms, and barnacles (Stanley et al. 2014). Sounds from fishing vessels may attract birds, seals, and dolphins, which then feed on the bait or catch (Söffker et al. 2015). This attraction to a food source elicited by anthropogenic noise is called the "dinner bell effect."

In terms of the potential negative effects of anthropogenic noise on animals, Fig. 13.1 shows a generalized view of increasingly severe effects closer to the noise source. Depending on where the noise source and the receiving animals are located in space, received noise will differ in spectral and temporal characteristics (see Chaps. 5 and 6 on sound propagation in air and water, respectively). While there are widely varying sound propagation conditions depending on the specific environment in which a sound is produced and received, received levels generally attenuate or decrease as sound propagates from its source. Given that no habitat is acoustically homogeneous or isotropic, received levels vary with azimuth (direction) and inclination (height or depth), leading to different impact ranges in all directions.

The absolute range and order of noise impact severity can differ based on features of the propagation environment, exposure context, and species involved (Ellison et al. 2012). In general, at the longest ranges, a noise might barely be audible to an animal and may be less likely to have any negative effect. Audibility of a noise depends on its amplitude and spectrum, propagation

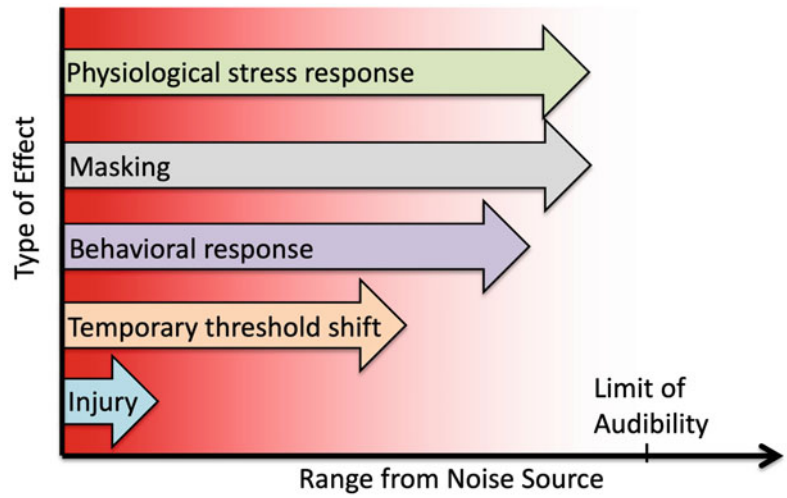



**Fig. 13.1** Sketch of generalized ranges from a noise source, at which different types of impact may occur

conditions from the source to the receiver, ambient noise conditions, and hearing abilities of the animal.

Stress is a physiological response, which might occur at long and short ranges and at low and high noise levels. Stress can be a direct response to noise (e.g., if a novel noise is suddenly heard) and an indirect response to noise (e.g., if masking causes stress). Stress can affect numerous life functions (including immune response, reproductive success, predator avoidance, etc.; Tarlow and Blumstein 2007).

Acoustic masking might occur over long ranges when a distant noise masks a faint signal. Masking is the process (and amount) by which the audibility threshold for a sound is raised by the presence of another sound (i.e., noise; American National Standards Institute 2013).[1] The higher the noise level is, the greater the masking effect. Masking can interfere with signals important to animals, such as their social communication calls, mother-offspring recognition sounds, echolocation signals, environmental sounds, or sounds by predators and prey (Dooling and Leek 2018). The animal's auditory system splits incoming sound into a series of overlapping bandpass filters, thus optimizing SNR in the bands occupied by the signal and enabling parallel processing (Moore 2013). The critical ratio is the most commonly measured parameter related to auditory masking. It is defined as the mean-square sound pressure of a narrowband signal (e.g., a tone) divided by the mean-square sound pressure spectral density of the masking noise at a level, where the signal is just detectable (see Chap. 10 on audiometry; International Organization for Standardization 2017). There are two categories of masking. Energetic masking occurs when the masking sound overlaps with the signal in both frequency and time, such that the signal is inaudible. Informational masking occurs later in the auditory process; the signal is still audible, but it cannot be disentangled from the masker (Moore 2013).

Somewhat closer to the source, changes in behavior of varying severity might be seen. An animal might change its orientation, cease prior behavior (e.g., feeding), move away from the source, or alter its vocal behavior, which may have implications for social functions.

Animals must be closer to sound sources to receive sound levels sufficiently high for noise-induced hearing loss (NIHL). NIHL results from overstimulation of the sensory cells in the inner ear, leading to metabolic exhaustion of the hair cells, damage to the organ of Corti, and in extreme cases, degeneration of retrograde

[1] ANSI/ASA S1.1 & S3.20 Standard Acoustical & Bioacoustical Terminology Database; https://asastandards.org/asa-standard-term-database/

ganglion cells and axons. NIHL includes both temporary and permanent loss of hearing, termed temporary threshold shift (TTS) and permanent threshold shift (PTS), respectively. Both TTS and PTS depend on the spectral and temporal (duration of exposure and duty cycle) characteristics of the noise received (Moore 2013; Saunders and Dooling 2018). TTS, by definition, is recoverable, but the time to recover depends on the amplitude, frequency, rise time, and duration of noise exposure. While experiencing TTS, animals could have a decreased ability to communicate, interact with offspring, assess their environment, detect predators or prey, etc. While TTS implies a full recovery without physical injury, TTS might still involve submicroscopic physical damage. Kujawa and Liberman (2009) showed that for high levels of TTS, sensory hair cells appear unharmed, yet afferent nerve terminals might be injured leading to cochlear nerve degeneration. Death of sensory hair cells in the ear, damage to the auditory nerve, or injury to tissues in the auditory pathway may lead to PTS (Liberman 2016).

At high levels of noise exposure, animals may incur injury (i.e., acoustic trauma) to tissues and organs, such as damage to ear bones, lungs, kidney, or gonads (Popper et al. 2014). In aquatic species, fast changes in pressure can cause blood gases to exit solution and gas-filled tissues or organs (e.g., swim bladders in fish) to expand and contract rapidly, which may damage surrounding tissues and organs (e.g., rupture the swim bladder). Rapid changes in sound pressure are more likely to cause damage than gradual changes (Popper et al. 2014).

Whether the effect of noise is auditory, behavioral, or physiological, individual animals of the same species or population respond at different ranges and in different ways. Age, health, sex, individual hearing abilities, prior experience (habituation versus sensitization), context, current behavioral state, and environmental conditions may all affect the responses of individuals. For example, bowhead whale (*Balaena mysticetus*) and gray whale (*Eschrichtius robustus*) responses to seismic surveys ranged from none-observed to moderate (i.e., changing vocalization rates and swimming behavior; Blackwell et al. 2015; Malme et al. 1983; Miller et al. 2005). Therefore, some studies have developed a dose-response curve (Fig. 13.2) relating likelihood of response (or percentage of a population that might respond) to the received level of the specific source of noise under consideration (e.g., Hawkins et al. 2014; Miller et al. 2014; Williams et al. 2014).

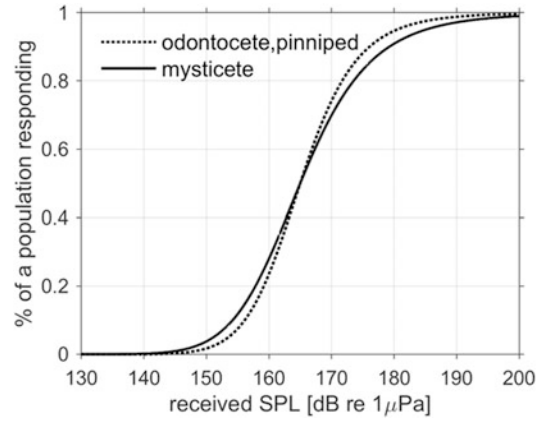



**Fig. 13.2** Example of a historical dose-response curve based on received exposure level as a metric of sound dose used to assess the likelihood of bioacoustic impact from mid-frequency sonar (Department of the Navy 2008). Half of a population was modeled to respond at 165 dB re 1 μPa, with fewer animals responding at lower levels, and more animals responding at higher levels

The effects of noise discussed so far, and the concepts of impact ranges (Fig. 13.1) and dose-response curves (Fig. 13.2) relate to acute noise exposures (e.g., to a single discharge of a seismic airgun array or a single supersonic overflight). The scientific difficulty is to link short-term, individual impacts to long-term, population-level impacts, considering that animals might travel and be exposed to aggregate noise from multiple sources distributed through space and time. While some studies have documented long-term reductions in species abundance and diversity (e.g., near highways or in industrialized areas; Francis et al. 2009; Goodwin and Shriver 2011), in the majority of cases (i.e., species and noise sources), it remains unknown how the impacts on individuals accumulate over time (i.e., over multiple exposures) and over a population.

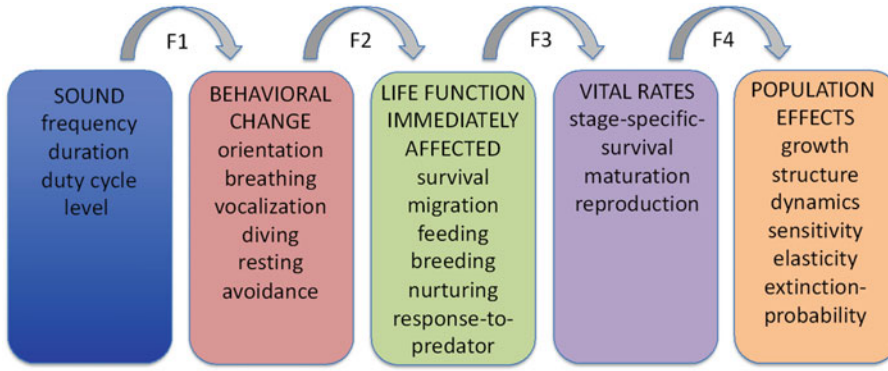



**Fig. 13.3** Population Consequences of Acoustic Disturbance (PCAD) model (National Research Council 2005), which links noise exposure from individual to population-level consequences via a series of stages, connected by transfer functions

Extrapolating temporary effects on individuals to population-level effects is problematic. The Population Consequences of Acoustic Disturbance (PCAD) model (Fig. 13.3) was originally developed for marine mammals and provides a framework for the link between noise exposure and population impacts (National Research Council 2005). The link is broken down into five stages and four transfer functions.

Data to fully parameterize this model are not available for any species. However, progress has been made for a few selected species, with the elephant seal (*Mirounga angustirostris*) being an excellent model in the marine world, having been studied extensively over long periods (Costa et al. 2016). This conceptual model has recently been more fully developed mathematically and broadened to consider potential changes in vital rates to estimate population-level effects of any form of disturbance (New et al. 2014); the resulting framework is now more broadly termed the Population Consequences of Disturbance (PCoD) model. Furthermore, novel conceptual paradigms have been proposed to consider population consequences of noise exposure from multiple stressors, complex interactions of which may be additive, synergistic, or antagonistic (Ocean Studies Board 2016). These models have implications for other taxa and their conservation management.

One important aspect of noise impact management is mitigation. To reduce the risk of impacts from acute noise exposure (e.g., from a marine seismic survey or detonation), the surrounding area is commonly observed (e.g., visually or acoustically), and operations are changed (e.g., temporarily reducing power or shutting down) if animals are detected within the so-called safety zones (Fig. 13.4; Weir and Dolman 2007). Sometimes, alternative (e.g., quieter) technology is available. Also, noise barriers may be employed (e.g., temporary, sound-absorbing walls in terrestrial environments, or bubble curtains in marine environments; Bohne et al. 2019). Operations may be ramped up in an attempt to warn animals (e.g., Wensveen et al. 2017). Short-term operations may be timed to avoid biologically critical seasons or habitats.

In the case of chronic noise, such as from shipping, voluntary area-wide speed reductions reduced noise levels (Joy et al. 2019). Similarly, voluntarily turning off engines in drive-through national parks is encouraged (Fig. 13.5). For long-term operations or installations (such as highways), permanent sound barriers are commonly erected in the terrestrial environment. But these mitigation measures can reduce habitat connectivity. Instead, overpasses and long underground roadways may shelter large areas from noise exposure while concurrently increasing habitat connectivity. Understanding the role sound plays in habitat fragmentation will increase the ability to make barriers, underpasses, and overpasses more effective at reducing noise exposure, while also increasing landscape connectivity.

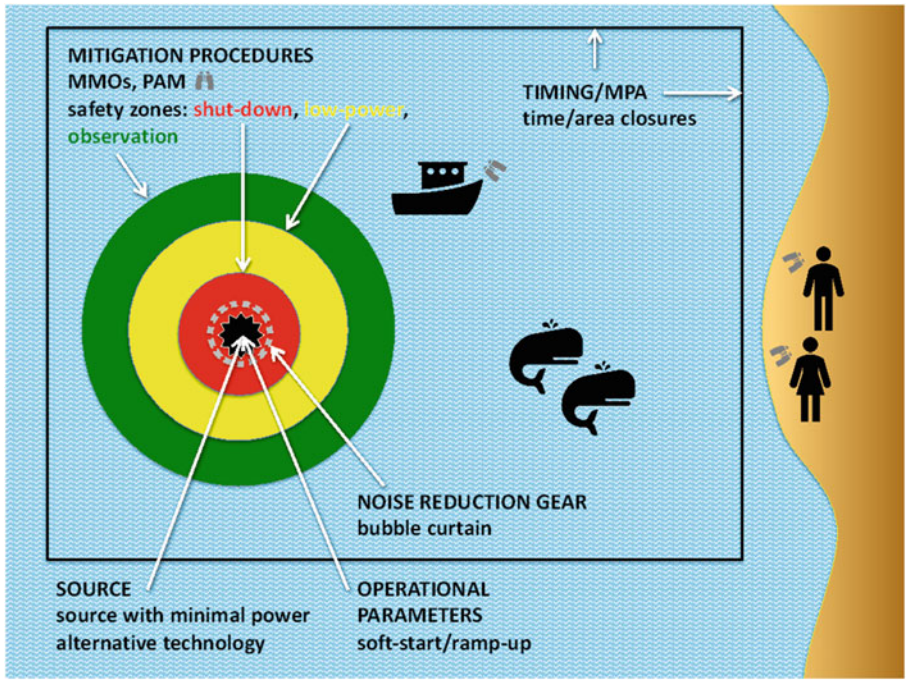



**Fig. 13.4** Bird's-eye sketch of different mitigation methods employed in the marine environment to reduce the risk of noise impacts (Erbe et al. 2018). The offshore, noise-producing platform is indicated by the black star. It is surrounded by safety zones, which are observed in real time. MMO: marine mammal observer, who might be on shore, or on the operations platform, or on an additional vessel. PAM: passive acoustic monitoring using hydrophones, possibly as a towed array. Operations temporarily reduce power or shut down if animals are detected within these zones and resume once animals have departed. In addition, modifications might be possible to the source or its operational parameters. Noise reduction gear (e.g., a bubble curtain around pile driving in shallow water) is indicated by gray dots. MPA: marine protected area, which might only be accessible during low-risk seasons

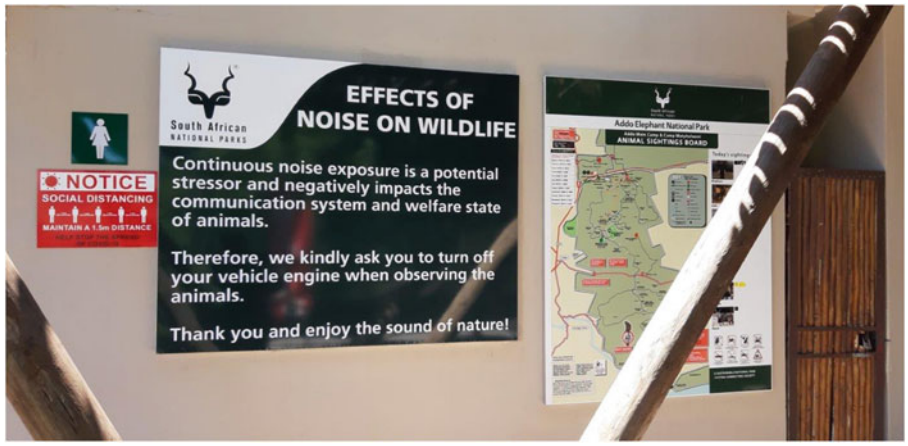



**Fig. 13.5** Photograph from Addo Elephant National Park, South Africa, encouraging visitors to switch off their car engines to limit noise effects on wildlife (courtesy of Cathy Dreyer, Conservation Manager, Addo Elephant National Park)

Overall, the effects of anthropogenic noise are a challenge to researchers, noise producers, and policy makers. Often, stakeholders have data from only a few studies on a few species from which to develop criteria for noise exposure. This chapter gives examples of the effects of noise on a variety of animal taxa.

## 13.2 Behavioral Options in a Noisy Environment

When exposed to anthropogenic noise, animals have choices of responses. Behavioral changes are perhaps the most frequently observed and reported effects of noise. In many cases, such changes might be an "affordable" adaptation, for example when an animal temporarily moves away from the noise. The response (or lack thereof) is likely based on a cost-benefit ratio or the cost of change to improve fitness versus the magnitude of the benefit by changing. Although a variety of behavioral changes in response to noise have been studied in several species, their implications for biological fitness are difficult to determine.

### 13.2.1 Habituation

Animals sometimes habituate to anthropogenic noise. Habituation is a form of learning in which an animal reduces or ceases its response to a stimulus after repeated presentations; in other words, the animal learns to stop responding to anthropogenic noise when it learns there are no significant consequences. Habituation can be difficult to determine in the wild. A lack of observed behavioral response does not necessarily mean that there was no response or that the animal habituated; the response might have been too small to be observed, or it was of physiological type, or the animal's hearing sensitivity might have been reduced by prior exposure.

There are many accounts of animals living without apparent detrimental impacts in areas of high ambient noise, for example small mammals that live and breed along runways, railroad tracks, or highways. The densities of white-footed mice (*Peromyscus leucopus)* and eastern chipmunks (*Tamias striatus*) did not decrease near roads. While both species were significantly less likely to cross a road than move the same distance away from roads, traffic volume (and noise level) had no effect (McGregor et al. 2008). Wale et al. (2013b) investigated the physiological responses of shore crabs (*Carcinus maenas*) to single and multiple ship-noise playbacks. Crabs consumed more oxygen, indicative of a higher metabolic rate and potential stress, when exposed to ship noise compared to ambient noise. However, repeated exposures to ship noise showed no change. The authors proposed that crabs exhibited the maximum response on the first exposure to ship noise, then habituated or became tolerant of the noise.

Even when no behavioral response is detectable, animals might accept noise exposure at levels that could have long-term hearing impacts, especially if there are benefits of sticking around. For example, each winter endangered manatees (*Trichechus manatus*) congregate around power plants in Florida likely in order to stay in the warm water effluence produced by the plant. In the process, they are potentially exposed to high levels of underwater noise for long periods. Seemingly, the benefit of the warm water outweighs the cost of noise exposure (JA Thomas, pers. obs.). Similarly, seals depredating at aquaculture sites might accept hearing loss inducing noise levels from acoustic harassment devices or "seal scarers" (Coram et al. 2014).

### 13.2.2 Change of Behavior

Temporary behavioral responses have been reported for gray whales that took a somewhat wider route around the noise from offshore oil drilling platforms, while continuing their normal round-trip migration from Alaska to Mexico (Malme et al. 1984). Such a subtle response likely won't have any long-term impact on fitness. Harbour porpoises (*Phocoena phocoena*), on the other hand, have been shown to forage almost

continuously around the clock and hence even moderate occurrences of anthropogenic disturbance might have significant fitness consequences (Wisniewska et al. 2016).

A permanent displacement from habitat has been suggested in egrets (*Ardea alba*) and great blue herons (*Ardea herodias*), judged by the altered distribution of nests along the Mississippi River, potentially in response to increased vessel traffic, such as tugboats and barges (JA Thomas, pers. obs.). A long-term displacement lasting six years occurred in killer whales (*Orcinus orca*) in response to acoustic harassment devices installed in parts of their habitat. Whales returned when devices were removed (Morton and Symonds 2002).

Noise affects not only animal movement but also other behaviors. Chaffinches (*Fringilla coelebs)* reduced their food pecking during increased background noise, which increased their vigilance; however, the increased alertness and hence reduction in predation risk might have reduced fitness via the reduction in food intake (Quinn et al. 2006). Similarly, California ground squirrels (*Otospermophilus beecheyi*) showed increased vigilance near wind turbines, potentially at the cost of other behaviors (Rabin et al. 2006). In the marine environment, anthropogenic noise interfered with the predator-prey relationship. Motorboat noise elevated metabolic rate in prey fish, which then responded less often and less rapidly to predation attempts. Predator fish consumed more than twice as much prey during boat noise exposure (Simpson et al. 2016).

Reinforcing an acoustic communication message with a visual display can enhance communication in a noisy environment. For example, male foot-flagging frogs (*Dendropsophus parviceps*) live in neotropical areas with fast-flowing streams, high levels of rain, and numerous other species of calling frogs. Foot-flagging frogs evolved the visual signal of stretching out one or two hind legs, vibrating their feet, or stretching out their toes while calling, assisting with their communication (Amézquita and Hödl 2004).

### 13.2.3 Change of Acoustic Signaling

Vocal behaviors can also change in response to noise. To reduce interference from urban daytime noise, chaffinches sang earlier in the day and European robins (*Erithacus rubecula*) changed vocal activities to nighttime (Bergen and Abs 1997; Fuller et al. 2007). The cost of this change in vocal behavior is unknown. Animals might also change the characteristics of their sounds to avoid masking. Changes in vocal effort such as increases in amplitude, repetition rate, and duration, or frequency shifts are collectively known as the Lombard effect, which has been demonstrated in several taxa, including frogs (Halfwerk et al. 2016), birds (Slabbekoorn and Peet 2003), and cetaceans (Scheifele et al. 2005). The Lombard effect has also been observed during odontocete echolocation: A captive beluga whale (*Delphinapterus leucas*) increased the amplitude and frequency of its echolocation signal when moved from a quiet habitat in San Diego to an area with high snapping shrimp noise in Hawaii (Au et al. 1985).

Some animal taxa might be limited in their ability to voluntarily and temporarily change the spectrographic features of their sounds—often called behavioral plasticity. Insects, for example, generate sound by stridulation of body parts, the resonance of which cannot be actively controlled. Consequently, a Lombard effect failed to be observed in *Oecanthus* tree crickets (Costello and Symes 2014); however, grasshoppers (*Chorthippus biguttulus*) from noisy habitats or those exposed to noise as nymphs produced higher-frequency sounds with higher duty cycles (i.e., increased sound-to-pause ratio), indicating developmental plasticity (Lampe et al. 2012, 2014).

A cessation of sound emission in the presence of anthropogenic noise can also occur. Thomas et al. (2016) studied the effects of construction noise on yellow-cheeked gibbons (*Nomascus gabriellae*) at Niabi Zoo. Before construction, a bonded pair and their four-year-old offspring

were quite soniferous. The pair commonly duetted in the early morning and displayed behaviors typical of a bonded pair. Once construction near their exhibit commenced, they gradually vocalized less often, and by the end of the four-month construction period, the pair bond had dissolved and the young became ill (possibly due to decreased quality of care with the loss of parent pair bond). For about a year, the pair remained distant from each other and did not vocalize. One of the authors (JA Thomas) played back recordings of the pair's own duet and those of wild gibbons. Already during the first playback, the pair slowly started to vocalize and move to the top of the exhibit where they normally performed their duet. They vocalized in response to their own duet as opposed to playbacks of other gibbon duets. The pair continued duetting for several more years of observation.

## 13.3 Physiological Effects

In addition to eliciting changes in fine- or gross-motor behavior and acoustic behavior, sound can also cause physiological impacts, like stress, hearing loss, or injury to tissues and organs. An animal with impaired hearing might exhibit different responses to sound and different acoustic behavior, compared to an animal with normal hearing.

A stress response may occur when noise is loud, novel, or unexpected (Wale et al. 2013a, b). Studies often concentrate on the effects of noise-induced stress on reproduction. However, stress also can result in: (1) a reduction or cessation of normal movement, with a reduced likelihood of escaping a predator; (2) reduced appetite, feeding, or food acquisition; and (3) excessive anti-predation behaviors. Attention is required to capture prey or avoid detection by a predator. Many animals use auditory cues to detect the presence of predators or prey, and any noise-induced distraction could limit this detection (Siemers and Schaub 2011). Chan et al. (2010) termed this the "distracted prey hypothesis".

The consequences of elevated stress levels can be far-reaching. Tarlow and Blumstein (2007) reviewed the effects of increased stress in birds resulting from human disturbances. The review documented changes in hormone levels, changes in heart rate, immunosuppression, changes in flight-initiation distance, disturbed breeding success, altered mate choice, and fluctuating anatomical asymmetry—all as a result of stress. While there have not been many long-term studies of noise-induced, chronic stress in animals, there is plenty of evidence from humans documenting, for example, hypertension and cardiovascular disease (Bolm-Audorff et al. 2020; Hahad et al. 2019; World Health Organization 2011).

Noise can further affect other non-acoustic sensing and information use (termed cross-modal impacts). For example, road noise impacted the ability of mongoose (*Helogale parvula*) to smell predator feces, leaving these mammals more susceptible to predation and loss of group cohesion (Morris-Drake et al. 2016). The effects of noise are complex and they differ by species. The following sections describe observed responses to sound by different taxa.

## 13.4 Noise Effects on Marine Invertebrates

Marine invertebrates comprise a great diversity of fauna with a corresponding diversity of sensory systems and modes of detecting sound or vibration. Only a few publications exist on the impacts of underwater sound on marine invertebrates.

### 13.4.1 Marine Invertebrate Hearing

Invertebrate species exhibit a diversity of sensory systems for detecting sound and vibration. Many crustaceans and molluscs have acoustic sensory systems that are an analogue to the fish otolith hearing system as they contain statocysts. These are small organs that house a dense mass (i.e., a statolith), which moves in response to sound and thus drives sensory hair cells, which create the

nervous response to the appropriate stimuli. Statocysts are involved in balance and motion sensing (e.g., in squids and cuttlefish; Arkhipkin and Bizikov 2000). Invertebrates can sense the particle motion of an incoming sound wave with the statocyst system, as reported, for example, in common prawn (*Palaemon serratus*; Lovell et al. 2005), octopus (*Octopus ocellatus*; Kaifu et al. 2008), and longfin squid (*Loligo pealeii;* Mooney et al. 2010).

Benthic molluscs, which are site-attached and fixed to the substrate, possess statocysts. These animals may be responsive to water-borne sound, to substrate-borne sound, or to sound waves traveling along the seabed-water interface. Some high-energy sound sources (e.g., impulsive seismic survey signals) can directly excite the ground (Day et al. 2016a). A benthic animal might derive information on nearby surf conditions or on an approaching predator grubbing along the seafloor from seabed-transmitted sound. Thus, benthic invertebrates, including molluscs and crustaceans, may be adapted to sense substrate-borne sound, as well as respond to water-borne sound.

Other invertebrates do not possess statocyst organs. Many invertebrates may be comprised primarily of soft tissue with no organs containing internal masses capable of exciting hair cells. Small animals of a single or few cells might merely vibrate in phase with the sound wave. Other vibratory sensory systems documented in invertebrates include single sensory hairs or antennal organs, such as in the copepod *Lepeophtheirus salmonis,* which responded to low-frequency vibrations or infrasound (<10 Hz; Heuch and Karlsen 1997).

Invertebrate larvae undergo multiple developmental stages of which the later stages, just before settlement, have the most developed sensory systems. These pre-settlement larvae are critical for recruitment success and thus of great concern with regard to anthropogenic impacts. Many late-stage larvae are responsive to sound cues for settlement; for example, those of corals (Vermeij et al. 2010) and crabs (Stanley et al. 2009). Information on the responses of late-stage larvae to anthropogenic sound is limited.

### 13.4.2 Effects of Noise by Taxon

Invertebrate statocyst systems can be overexcited by excessive motion of the statolith in response to intense sound, resulting in damage to surrounding hair cells or membranes, as observed in lobsters exposed to seismic airguns (Day et al. 2016a, 2019). There were no signs of repair over the 365-day holding period in these lobsters. While such damage likely results in a degradation of an animal's sensory capability, the degree to which the fitness of wild animals is affected remains unclear and in at least one documented case did not seem to alter population success (Day et al. 2020).

Invertebrates comprised of soft tissue with no dense masses might vibrate with a sound wave. In the case of intense impulse signals, this mechanical motion might cause physiological trauma to cells, although the onset level is not known (Lee-Dadswell 2011). Planktonic invertebrates with no statocyst systems but with sensory appendages and antennal organs have been shown to be susceptible to damage from intense impulse signals (McCauley et al. 2017).

Studies on noise effects on marine invertebrates show a range of impacts from none to severe, and results are difficult to compare due to vastly different experimental regimes. The following sections provide examples of study results on a species level.

#### 13.4.2.1 Squid

Caged squid (*Sepioteuthis australis*) that were approached by a 20-in$^3$ airgun moved away from the airgun at received sound exposure levels (SEL) of 140–150 dB re 1 μPa$^2$s and spent more time near the sea surface; a strong startle response of the squid inking and jetting away from the airgun was observed when the airgun was discharged at about 30-m range with a received SEL of 163 dB re 1 μPa$^2$s (Fewtrell and McCauley 2012; McCauley et al. 2003a). Two events of giant squid (*Architeuthis dux*) mass mortality in the Bay of Biscay in 2001 and 2003 were suggested to have been a result of marine seismic surveys, based on tissue damage (Guerra

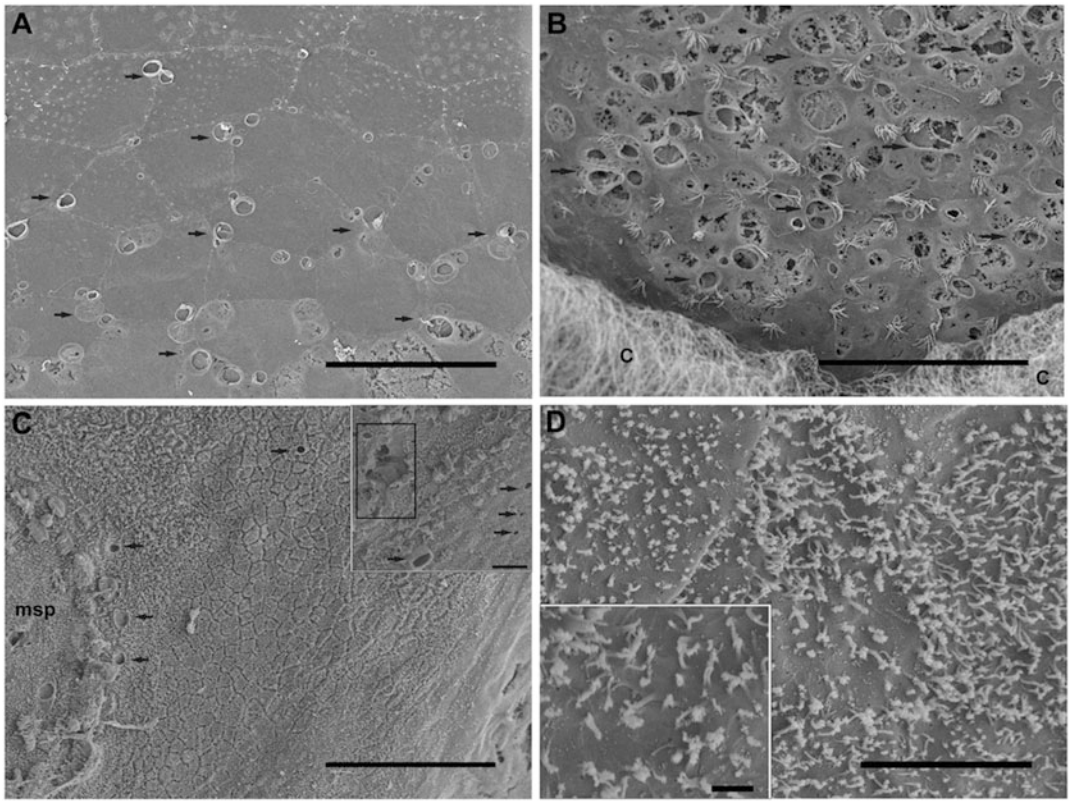



**Fig. 13.6** Scanning electron microscope images of squid (*Illex coindetii*) epithelium 48 h after sound exposure. Arrows point to missing cilia and holes. Scale bars: A, B, C = 50 μm, D = 10 μm (Solé et al. 2013). © Solé et al.; https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0078825; licensed under CC BY 4.0; https://creativecommons.org/licenses/by/4.0/

et al. 2004). Statocyst hair cell damage was found in cephalopods (cuttlefish and squid) subjected to simulated sonar sweeps in a laboratory tank (André et al. 2011; Solé et al. 2013; Fig. 13.6).

#### 13.4.2.2 Scallops

Scallops (*Pecten fumatus*) exhibited behavioral changes as a result of exposure to a 150-in$^3$ airgun, which continued during the full 120-day post-exposure monitoring, suggesting damage to the statocyst organ, which controls balance (Day et al. 2016a, 2017). Physiological measures changed for the worse and mortality increased with dose from 1 to 4 passes of the airgun (Day et al. 2016a, 2017). A different study failed to find any significant effects of seismic airguns on scallops (Parry et al. 2002); however, animals had been removed from their seafloor habitat and were suspended in lantern nets in the water column where they would not have experienced substrate-borne and interface (i.e., at the seafloor) sound and vibration. Also, physiological measurements and long-term monitoring were not conducted. Przeslawski et al. (2018) made observations of wild scallops exposed to seismic airguns and found no discernible impacts, but the study had insufficient controls and no physiological measurements, and longer-term post-exposure sampling was not undertaken.

#### 13.4.2.3 Crustaceans

Spiny lobsters (*Jasus edwardsii*) were exposed to single passes of a 45 or 150-in$^3$ airgun and monitored for 365 days after exposure (Day et al.

2016a). No mortality or significant morphological changes were found in adults or in egg viability (Day et al. 2016b). However, impaired righting ability correlating with damaged statocyst organs (ablated hair cells) and compromised immune function were reported (Day et al. 2019; Fitzgibbon et al. 2017). How these changes would impact wild lobsters is unclear, especially as another study using an apparently healthy lobster population found pre-existing statocyst damage and no further increase in damage after experimental airgun exposure, suggesting the animals had been exposed to intense noise *in situ* before the experiment but had adapted to the damage (Day et al. 2020). American lobsters (*Homarus americanus*) exposed to 202–227 dB re 1 μPa pk-pk airgun signals in a large tank exhibited physiological changes but no impact on righting times and no mortality (Payne et al. 2007). Andriguetto-Filho et al. (2005) compared shrimp (*Litopenaeus schmitti, Farfantepenaeus subtilis,* and *Xyphopenaeus kroyeri*) catch rates before and after airgun exposure (635 $in^3$) in shallow (2–15 m) water in north-eastern Brazil, finding no difference. The playback of ship noise as opposed to ambient noise negatively affected the foraging and antipredator behavior of shore crabs (*Carcinus maenas;* Wale et al. 2013a). Furthermore, oxygen consumption was greater during ship noise playback (possibly a stress response), and heavier crabs were more affected (Wale et al. 2013b). Evidently, there might be different responses to anthropogenic noise, depending on the size of an individual organism.

#### 13.4.2.4 Coral

Experiments on the potential impacts of a 2055-$in^3$ 3D seismic survey on corals were undertaken in the 60-m deep lagoon of Scott Reef, north-western Australia. Corals within and outside of the lagoon were exposed to airgun noise over a 59-day period. Some corals received airgun pulses from straight overhead (seismic source at 7-m depth, corals at ~60-m depth), whereas the full seismic survey passed within tens to hundreds of meters horizontal offset, yielding maximum received levels of 226–232 dB re 1 μPa pk-pk, 197–203 dB re 1 $\mu Pa^2s$, and 214–220 dB re 1 μPa rms (McCauley 2014). No evidence of mechanical trauma (i.e., breakage), physiological impairment (i.e., polyp withdrawal or reduction in soft coral rigidity), or long-term change in coral community structure was found (Battershill et al. 2008; Heyward et al. 2018).

#### 13.4.2.5 Larvae/Plankton

Noise and vibration from ships can enhance the settlement and growth of larvae of bryozoans, oysters, calcareous tubeworms, and barnacles, and thus increase biofouling (Stanley et al. 2014). The effects of a 150-$in^3$ airgun were studied by Day et al. (2016b) with berried (with eggs) spiny lobster (*Jasus edwardsii*) off Tasmania. No mortality of adult lobster or eggs could be attributed to the airgun at cumulative received SEL of up to 199 dB re 1 $\mu Pa^2s$. Some differences in exposed larvae morphology were noted (i.e., slightly larger than controls), but no differences in larval hatching rates or viability were found. These were early-stage larvae with underdeveloped sensory organs; results might differ for late-stage larvae. Parry et al. (2002) found no impacts on plankton from a 3542-$in^3$ seismic array, but their statistical power to detect impacts was low. Aguilar de Soto et al. (2013) exposed early-stage scallop larvae to airgun signals simulated by an underwater loudspeaker 9 cm away from the larval tank. Morphological deformities were found in all exposed larvae. However, the exact stimulus was unknown owing to the experimental setup and inherent acoustic limitations in small tanks.

McCauley et al. (2017) reported negative impacts, including a 2–3 times greater mortality rate, on various zooplankton out to 1 km from passage of a 150-$in^3$ seismic airgun. In contrast, Fields et al. (2019) exposed constrained adult North Sea copepods (*Calanus finmarchicus*) to a 520-$in^3$ airgun cluster with measured impacts limited to within 10 m. McCauley et al. stated that the "'copepods dead' category was dominated by the smaller copepod species (*Acartia tranteri*, *Oithona* spp.)". These species are ~0.5 mm in length as compared to the ~2.5-mm *C. finmarchicus*, suggesting a possible size dependency for impacts from airguns. The 1-km

impact range given by McCauley et al. (2017) was within the repeat range (400–800 m) within which a 3D seismic survey vessel would pass on an adjacent seismic line, so that the entire survey area could have its plankton field degraded. Richardson et al. (2017) ran ecological models to assess the scale of this impact. Assuming an area of strong tidal currents and consistent ocean current, a 3-day copepod turnover rate, and a three-fold increase in copepod mortality within 1.2 km, the copepod plankton field was modeled to recover within three days of completion of a mid-size 3D seismic survey. But, when Richardson et al. (2017) reduced the strength of the currents in the model, the impact persisted for three weeks. Many larger zooplankton have a longer than 3-day turnover rate (i.e., weeks to months) with larval forms having a once or twice per year recruitment cycle, enhancing impacts above the published model output. Given the central role zooplankton play in the ocean ecosystem, and given that not all turn over rapidly, the results of McCauley et al. (2017) are of concern for ocean health.

## 13.5 Noise Effects on Terrestrial Invertebrates

Soniferous terrestrial invertebrates include some crabs, spiders, and insects. Limited information exists on the impacts of sound on terrestrial invertebrates, with insects being the main group studied. Currently, little is known about how egg and larvae of terrestrial invertebrates respond to high-amplitude anthropogenic sounds. As a result, this section concentrates on adult insects as representatives of terrestrial invertebrates.

### 13.5.1 Insect Hearing

The ability to hear air-borne sound evolved independently at least 24 times in seven orders of insects (Greenfield 2016), either as tympanal hearing or hearing with antennae. These ears are sensitive to a very broad range of frequencies, from less than 1 kHz to high ultrasonics beyond 100 kHz. Signaling at these frequencies is important for mate attraction and localization, rivalry, and spacing of individuals within populations. In addition, many species use their ears to detect and avoid predators. Some species of flies eavesdrop on calling insects to locate and parasitize them.

An evolutionary adaptation to ambient noise from competing insect choruses is the modification of peripheral sensory filters, such as the sharpening of tuning in the cricket (Fig. 13.7). Such sharp tuning curves reduce the amount of masking noise within the filter (Schmidt et al. 2011).

However, the most prevalent form of insect communication involves substrate-borne sound. More than 139,000 described taxa are expected to exclusively use vibrational signaling and an additional 56,000 taxa use a combination of vibrational communication and other forms of mechanical signaling (Cocroft and Rodríguez 2005). The sensory organs monitoring substrate-borne sound (e.g., the subgenual organs in the legs) are tuned to frequencies below 1 kHz and are extremely sensitive.

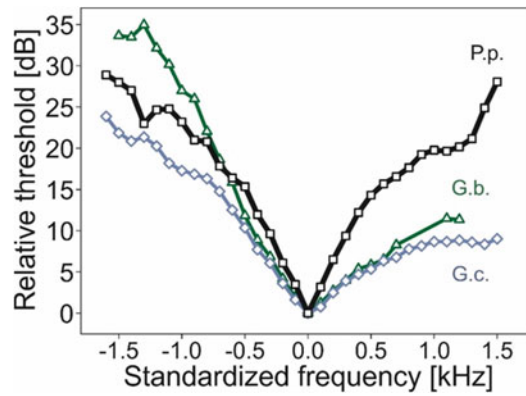



**Fig. 13.7** Graph of standardized mean sensitivity tuning curves of auditory interneuron AN1 in three cricket species: *Paroecanthus podagrosus* (P.p.), a neotropical cricket communicating under strong background noise levels, and *Gryllus bimaculatus* (G.b.) and *G. campestris* (G.c.), field crickets in environments with less background noise. The increased steepness in tuning toward higher frequencies filters out competing frequencies from other crickets (Schmidt et al. 2011). © Schmidt et al.; https://jeb.biologists.org/content/214/10/1754. Published green open access; https://jeb.biologists.org/content/rights-permissions

Anthropogenic noise sources produce significant amplitudes of air-borne sound at frequencies from less than 10 Hz to 50 kHz (e.g., traffic on roads and railways, compressors, wind turbines, military activities, and urban environments). At the same time, airport, road, and railroad traffic and construction are significant sources of low-frequency, substrate-borne vibrations below 1 kHz. Such substrate-borne noise may be created directly by vibrating the substrate (e.g., by driving over it) or indirectly via air-borne noise that induces vibrations in the substrate. The relatively low-frequency sound produced by many of these sources suffers less attenuation and can thus travel farther from the source. Because many insects have very sensitive receptors for substrate-borne sound, with displacement thresholds less than 1 nm, they are likely to detect anthropogenic sources over long distances. Anthropogenic noise may therefore have a significant impact on the ability of insects to communicate and listen in both the air-borne and substrate-borne channel (reviewed by Morley et al. 2014; Raboin and Elias 2019).

### 13.5.2 Behavioral Effects

Anthropogenic noise may impact insects in various ways. It can mask communication signals, increase stress, affect larval development, and ultimately decrease lifespan (reviewed by Raboin and Elias 2019). The most common consequence of noise is masking, when noise overlaps in time and frequency with a signal. This decreases the signal-to-noise ratio and thus the detection and/or discrimination of signals. For example, Schmidt et al. (2014) found that anthropogenic noise resulted in less effective female cricket orientation toward signaling males (phonotaxis: orientated movement in relation to a sound source), which, in crickets, is the usual way to bring the sexes together. In another cricket species, males shortened their calls and paused singing with increasing noise level. However, males did not adjust the duration of intervals between song elements important for species identification (Orci et al. 2016). Apparently, these insects can neither modify the fundamental frequency of their song nor increase the amplitude of their calls in noise (i.e., lack of a Lombard effect), as do some species of frogs and birds, to reduce masking by anthropogenic noise.

For insects using substrate-borne signals, experimentally induced noise may disrupt mating. Insects either respond less frequently to signals of the opposite sex, or they cease signaling during the initial part of communication (Polajnar and Čokl 2008). The fact that noise can disrupt substrate-borne communication between the sexes may be utilized in pest control in agriculture (Polajnar et al. 2015). For example, substrate-borne noise can mask the mating signals of species of leafhoppers, which represent a major pest in vineyards, resulting in reduced reproductive success. A similar approach was successful with pine bark beetles, when the substrate-borne noise spectrally overlapped with beetle signals (Hofstetter et al. 2014).

The failure to adjust the frequency or amplitude of mating signals in noise does, however, not exclude other means of behavioral plasticity. For example, the responses of male field crickets (*Gryllus bimaculatus*) to traffic noise depended on prior experience (Gallego-Abenza et al. 2019). Recordings of car noise were played back to males living at different ranges from the road and, therefore, with different prior experience to road noise. Males farther from the road decreased their chirp rate more than those nearer by, suggesting that "behavioral plasticity modulated by experience may thus allow some insect species to cope with human-induced environmental stressors" (Gallego-Abenza et al. 2019).

Developmental plasticity may also manifest in signal modifications in response to noise. The courtship signals of grasshoppers are more broadband in frequency than those of crickets. Specifically, male grasshoppers (*Chorthippus biguttulus*) from roadside habitats produced higher-frequency signals compared to grasshoppers in quieter habitats (Lampe et al. 2014). In an experiment that reared half of the grasshopper nymphs in a noisy environment and the other half in a quiet environment, adult males from the first group produced signals with higher-

frequency components, suggesting that developmental plasticity allows signal modifications in noisy habitats.

### 13.5.3 Physiological Effects

Strong anthropogenic noise can result in hearing loss. Auditory receptors in the locust ear showed a decreased ability to encode sound after noise exposure. The mechanism for such hearing loss reveals striking parallels with that of the mammalian auditory system (Warren et al. 2020). A series of experiments was conducted to determine whether exposure to simulated road traffic noise induces increased heart rates, as an indicator of a stress response (Davis et al. 2018). Larvae of the monarch butterfly (*Danaus plexippus*) exposed for 2 h to road traffic noise experienced a significant increase in heart rate, indicative of stress. Because these larvae do not have ears for air-borne sound, the likely sensory pathway involved vibration receptors. However, exposing larvae for longer periods (up to 12 days) to continuous traffic noise did not increase heart rate at the end of larval development; so chronic noise exposure may result in habituation or desensitization. However, habituation to stress during larval stages may impair reactions to stressors in adult insects.

While more research is necessary to understand the sensory strategies for avoiding or compensating for anthropogenic noise, there are some cases where insects experience a significant fitness advantage. This may happen in a predator-prey or parasitoid-host relationship, when the noise decreases the ability of a parasitoid fly to localize calls of their host crickets (Lee and Mason 2017), or when bats as predators of flying insects are less efficient foragers in the presence of anthropogenic noise (Siemers and Schaub 2011).

## 13.6 Noise Effects on Reptiles

Reptiles have both aquatic (sea turtles, alligators, and crocodiles) and terrestrial (geckos, snakes, iguana, whiptails, geckos, chameleons, gila monsters, monitors, and bearded dragons) species. Soniferous reptiles include some snakes, alligators, crocodiles, geckos, and freshwater and marine turtles (e.g., Young 1997).

Reptiles are surrounded by anthropogenic noise from traffic (in water, on land, and in air), construction, mineral and hydrocarbon exploration and production, etc. Because many anthropogenic noise sources are low in frequency and thus within the reptilian hearing range, understanding the impact of these sources on behavior and physiology is an important start for reptile conservation.

Little literature exists on the impacts of anthropogenic noise on reptiles, with sea turtles having received recent attention. Simmons and Narins recently reviewed the topic (2018). Currently, little is known about how eggs and juvenile reptiles respond to anthropogenic noise. As a result, this section concentrates on adult sea turtles as a representative of reptiles.

Acoustic signals play an important role in turtle social behavior and reproduction. Turtles make very-low-frequency calls of short duration by swallowing or by forcibly expelling air from their lungs. Galeotti et al. (2005) published a summary of sound occurrence, context, and usage in Cryptodira chelonians—a taxon, which is quite soniferous. In general, turtles call when mating or seeking a mate, when they are sick or in distress, or for other reasons. Male red-footed tortoises (*Chelonoidis carbonaria*) make a clucking sound during mounting, Greek tortoises (*Testudo graeca*) whistle during combat, and young big-headed turtles (*Platysternon megacephalum*) squeal when disturbed (Galeotti et al. 2005). Nesting female leatherback sea turtles (*Dermochelys coriacea*) make a belching sound (Cook and Forrest 2005; Mrosovsky 1972), and the sounds from leatherback sea turtle eggs are believed to help coordinate hatching (Ferrara et al. 2014).

### 13.6.1 Reptile Hearing

Not all reptiles produce sound for communication. Most reptiles can detect substrate-borne vibrations (e.g., Barnett et al. 1999; Christensen

et al. 2012). The auditory anatomy of most reptile species includes a tympanic membrane near the rear of the head, a middle ear with a stapes, and a fluid-filled inner ear housing the lagena and its sound-sensing cells (Wever 1978). Brittan-Powell et al. (2010) indicated that reptile hearing is similar in frequency range to hearing in birds and amphibians. The most sensitive lizards have similar absolute sensitivities to birds. Ridgway et al. (1969) used electrophysiological methods to test hearing abilities of the green sea turtle (*Chelonia mydas)* and found peak sensitivity between 300 and 400 Hz, with the best hearing range from 60 to 1000 Hz. In general, the best frequency range of hearing in chelonids (turtles, tortoises, and terrapins) is 50–1500 Hz (Popper et al. 2014).

### 13.6.2 Behavioral Responses to Noise

Sea turtles may be exposed to acute and chronic noise. The soundscape of the Peconic Bay Estuary, Long Island, NY, USA, a major coastal foraging area for juvenile sea turtles, was recorded during sea turtle season. There was considerable boating and recreational activity, especially between early July and early September. Samuel et al. (2005) suggested that increasing and chronic exposure to high levels of anthropogenic noise could affect sea turtle behavior and ecology. Indeed, loggerhead sea turtles have been shown to dive when exposed to seismic airgun noise—perhaps as a means of avoidance (DeRuiter and Larbi Doukara 2012). In the terrestrial world, desert tortoises (*Gopherus agassizii*) exposed to simulated jet overflights did not show a startle response or increased heart rate, but they froze; and in response to simulated sonic booms, they exhibited brief periods of alertness (Bowles et al. 1999).

Unfortunately, there is a complete lack of data on masking of biologically important signals in sea turtles and other reptiles by anthropogenic noise (Popper et al. 2014). Similarly, there has been little research on physiological effects of noise in reptiles.

## 13.7 Noise Effects on Amphibians

Frogs rely heavily on acoustic communication for mating. Noise has been shown to alter both the production and perception of frog vocalizations. This can have serious implications for reproduction in these animals. Males that do not call as often will not attract females to their locations along a pond edge. Females that do not hear the advertisement calls from the males will not be able to localize or approach them. Further, they will not be able to sample multiple males for selection of the most attractive one. Studies have been conducted in both the laboratory and the field to determine the effects of noise on acoustic communication in frogs, for both vocal production and auditory perception.

### 13.7.1 Frog Hearing

The amphibian ear consists of a tympanic membrane on the outside through which sound enters the ear, a middle ear containing a columella, similar to the mammalian stapes, that provides mechanical lever action, and an inner ear in which sound is converted to neural signals (Wever 1985). The inner ear contains two papillae, known as the amphibian papilla, which responds to lower frequencies, and the basilar papilla, which responds to higher frequencies. Audiograms show good sensitivity between 100 Hz and a few kHz (e.g., Megela-Simmons et al. 1985). Some species, however, exhibit sensitivity also to ultrasound (Narins et al. 2014), and others to infrasound (Lewis and Narins 1985).

### 13.7.2 Behavioral Responses to Noise

Some species of frogs, like other animals, are known to avoid roads and highways, possibly to avoid both traffic mortality and a reduced transmission of vocal signals (reviewed by Cunnington and Fahrig 2010). Several studies, however, failed to document behavioral avoidance of noise by frogs or did not find reduced

frog abundance near continuous noise sources such as highways (Herrera-Montes and Aide 2011).

Nonetheless, noise does affect the perception of acoustic signals by frogs. Bee and Swanson (2007) investigated the potential of noise from road traffic to interfere with the perception of male gray treefrog (*Dryophytes chrysoscelis*) signals by females. Using a phonotaxis assay, they presented females with a male advertisement call at various signal levels (37–85 dB re 20 μPa) in three masking conditions: (1) no masking noise, (2) a moderately dense breeding chorus, and (3) road traffic noise recorded in wetlands near major roads. In both the chorus and traffic noise maskers, female response latency increased, orientation behavior toward the signal decreased, and response thresholds increased by about 20–25 dB. The authors concluded that realistic levels of traffic noise could limit the active space, or the maximum transmission distance, of male treefrog advertisement calls. Another treefrog (*Dendropsophus ebraccatus*) tested in a laboratory to compare the effects of dominant frequency and signal-to-noise ratio on call perception showed a low-frequency call preference in quiet conditions (usually correlated with larger, more attractive males), but no preference at higher signal-to-noise ratios (Wollerman and Wiley 2002). These results indicate that females listening to males in a noisy environment will likely make errors in mate choice.

Sun and Narins (2005) examined the effects of fly-by noise from airplanes and played back low-frequency sound from motorcycles to an assemblage of frog species in Thailand. Three of the most acoustically active species (*Microhyla butleri, Sylvirana nigrovittata,* and *Kaloula pulchra*) decreased their calling rate and the overall intensity of the assemblage calls decreased. However, calls from another frog (*Hylarana taipehensis*) seemed to persist. The authors suggested that the anthropogenic noise suppressed the calling rate of some species, but seemed to stimulate calling behavior in *H. taipenhensis*. Another study found that the vocalization rate of European treefrog (*Hyla arborea*) decreased in traffic noise (Lengagne 2008). Barber et al. (2010) believed that these frogs were unable to adjust the frequency or duration of their calls to increase signal transmission. Penna et al. (2005) found a similar decrease in call rate in leptodactylid frogs (*Eupsophus calcaratus*) exposed to recordings of natural noise in the wild.

An effective way to increase the likelihood that acoustic signals will be received is by increasing the intensity of those signals (Lombard effect). Love and Bee (2010) measured the intensities of vocalizations produced in the laboratory by Cope's gray treefrog (*Dryophytes chrysoscelis*) in the midst of different levels of background noise, similar to a frog chorus. They found no evidence for the existence of the Lombard effect in their frogs. Frogs produced calls at a level of 92–93 dB re 20 μPa, regardless of noise level. Similar to findings from other frogs, Cope's gray treefrogs increased call duration and decreased call rate with increasing noise levels. However, they appeared to be maximizing their call amplitudes in every calling situation, which does not allow them to increase their call intensities further when needed. On the contrary, túngara frogs (*Engystomops pustulosus*) and rhacophorid treefrogs (*Kurixalus chaseni*) did increase their call levels in noise (Halfwerk et al. 2016; Yi and Sheridan 2019).

Another possible way for a frog to increase communication efficacy would be to increase the frequencies of their calls to be above the frequency of the masking noise. Parris et al. (2009) found that two species of frogs (southern brown treefrog, *Litoria ewingii*, and common eastern froglet, *Crinia signifera*) called at a higher frequency in traffic noise (e.g., 4.1 Hz/dB for *L. ewingii*), and suggested this was an adaptation to be heard over the noisy environmental conditions. An extreme form of this frequency-increasing behavior has been discovered in concave-eared torrent frogs (*Odorrana tormota*) in China (Feng and Narins 2008). These frogs live near extremely loud streams and waterfalls (58–76 dB re 20 μPa, up to 16 kHz), which should make vocalizations difficult for other frogs to hear, at least at the lowest frequencies. The calls from these frogs are quite different from the

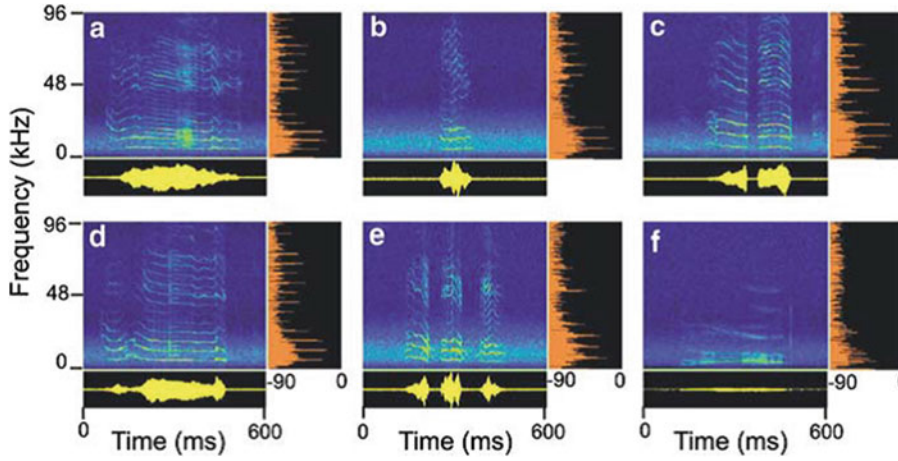



**Fig. 13.8** Spectrograms, waveforms, and call spectra from six vocalizations from the *O. tormota* frog (Feng and Narins 2008). Reprinted by permission from Springer Nature. A. S. Feng and Narins, P. M. Ultrasonic communication in concave-eared torrent frogs (*Amolops tormotus*). Journal of Comparative Physiology A, 194(2), 159–167; https://link.springer.com/article/10.1007/s00359-007-0267-1. © Springer Nature, 2008. All rights reserved

vocalizations of other frogs, however. These torrent frogs produce numerous vocalizations with energy in the ultrasonic frequency range (Fig. 13.8). A phonotaxis study found that female torrent frogs actually preferred synthetic male calls embedded in higher-amplitude stream noise than those embedded in lower-amplitude stream noise (Zhao et al. 2017). These ultrasonic signals are both produced and perceived by males and females, suggesting that they are not just a by-product of vocal production, and are instead an adaptation to avoid signal masking in a very noisy environment (Shen et al. 2008).

Some species of frogs are known to use visual signals when conditions are noisy, in an effort to improve communication. Grafe et al. (2012) recorded acoustic and visual communication strategies in noisy conditions by the Bornean rock frog (*Staurois parvus*). These frogs modified the amplitude, frequency, repetition rate, and duration of their calls in response to noise, but in addition engaged in visual foot-flagging and foot-touching behaviors. In a noisy world and with limited flexibility in vocal production capabilities, adding a visual component to an acoustic signal may be one of the only ways these animals are able to adapt.

### 13.7.3 Physiological Responses to Noise

Spatially separating a signal from a masker is one way to improve signal detectability. Spatial release from masking has been demonstrated in frogs behaviorally as well as physiologically. Ratnam and Feng (1998) recorded from single units in the inferior colliculus of northern leopard frogs (*Lithobates pipiens*) and found improvements in signal detection thresholds with spatially separated signals and noise maskers relative to spatially coincident signals and maskers. This has been shown in laboratory studies with awake behaving animals, when female Cope's gray treefrogs approached a target signal (male calling frog) more readily when it was spatially separated (by 90°) from a noise source (Bee 2007). This spatial release from masking, in the range of 6–12 dB, is similar to what is seen in other animals such as budgerigars (*Melopsittacus undulatus*; Dent et al. 1997) and killer whales (Bain and Dahlheim 1994).

Finally, increased levels of corticosterone, which correlated with impaired female mobility, have been shown in high traffic noise conditions in female wood frogs (*Lithobates sylvaticus*)

(Tennessen et al. 2014), although a recent study suggests that eggs taken from high traffic noise conditions yielded frogs that were less affected by noise exposure than frogs from eggs taken from low traffic noise environments, suggesting adaptations are possible (Tennessen et al. 2018). Whether it is from the stress or the masking of the acoustic signals, anthropogenic noise has been shown to have negative consequences.

## 13.8 Noise Effects on Fish

All fish species studied to date can detect sound. Hundreds of species are known to emit sound with the most prominent display of sound production in fishes being their choruses on spawning grounds (Slabbekoorn et al. 2010). Adult, juvenile, and larval-stage fishes actively use environmental sound to orientate and settle (Jeffrey et al. 2002; Simpson et al. 2005, 2007). Herring (*Clupea harengus*) have shown avoidance behavior to playbacks of sounds of killer whales, one of their predators (Doksaeter et al. 2009). Underwater anthropogenic noise can have a variety of effects on fish, ranging from behavioral changes, masking, stress, and temporary threshold shifts, to tissue and organ damage, and death in extreme cases (Hawkins and Popper 2018; Normandeau Associates 2012; Popper and Hastings 2009). Mortality can also result from an increased risk of predation in noisy environments (Simpson et al. 2016). Despite the growing amount of literature, our understanding of the cumulative effects of multiple exposures and the fitness implications to wild fish is limited.

### 13.8.1 Fish Hearing

Fish have two systems detecting sound and vibration: the inner ear and the lateral line system. The inner ear of fish resembles an accelerometer. It contains otoliths, which are bones of approximately three times the water density. Waterborne acoustic waves therefore result in differential motion between the otoliths and the fish's body, thus bending hair cells coupled to the otoliths of the inner ear, which sends neural signals to the brain. The inner ear is sensitive to particle motion. Fish with swim bladders close to or even connected to the ears are also sensitive to acoustic pressure. This is because the sound pressure excites the gas bladder, which reradiates an acoustic wave that drives the otolith. Particle motion then creates differential movement between the otoliths and the rest of the ear. The lateral line system involves neuromasts that detect water flow and acoustic particle motion. Due to variability in otolith anatomy and the absence or presence and variable connectivity of swim bladders, fish hearing varies greatly with species in terms of sensitivity and bandwidth, with most species sensitive to somewhere between 30 and 1000 Hz, but some species detecting infrasound, and others ultrasound up to 180 kHz (Popper and Fay 1993, 2011; Tavolga 1976). Hearing in noise has been studied and parameters such as the critical ratio (signal-to-noise ratio for sound detection, see Chap. 10) have been measured (Fay and Popper 2012; Tavolga et al. 2012); however, the significance of acoustic masking to fish fitness and survival remains poorly understood.

### 13.8.2 Behavioral Responses to Noise

The schooling behavior of fish has been observed to change in response to an approaching airgun with fish swimming faster, deeper in the water column, and in tighter schools (Davidsen et al. 2019; Fewtrell and McCauley 2012; Neo et al. 2015; Pearson et al. 1992). Caged fish had compacted near the center of the cage floor at received levels of 145–150 dB re 1 $\mu Pa^2s$ and swimming behavior returned to normal after 11–31 min (Fewtrell and McCauley 2012). A startle response was noted when the airgun was discharged at close range (Pearson et al. 1992), but not when the received level was ramped up by approaching from a longer range; also, the startle response diminished over time (Fewtrell and McCauley 2012). Wild pelagic and mesopelagic species dove deeper and their abundance increased at long range from the airgun array (Slotte et al. 2004). There are a few studies

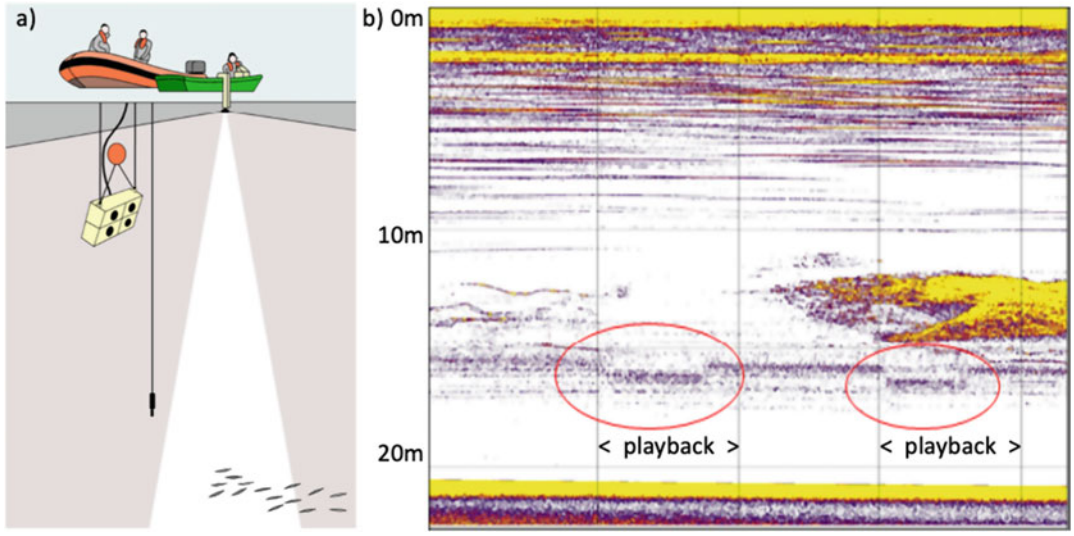



**Fig. 13.9** (**a**) Experimental setup to study fish responses to playbacks of pile driving sound. (**b**) Echogram of zooplankton dropping in depth below sea surface during playback of pile driving sound (red ellipses). Time is along the x-axis; playback started at the 1st vertical black line, stopped at the 2nd line, restarted at the 3rd line, and stopped at the 4th line (modified from Hawkins et al. 2014). © Acoustical Society of America, 2014. All rights reserved

documenting a drop in catch rates of pelagic fish after seismic surveying (Engas and Løkkeborg 2002; Engås et al. 1996; Slotte et al. 2004), believed to be due to behavioral responses.

Hawkins et al. (2014) played pile driving noise to wild zooplankton and fish. A loudspeaker was deployed from one boat for sound transmission, while an echosounder and side-scan sonar were deployed from a second boat for animal observation (Fig. 13.9a). Zooplankton dropped in depth below the sea surface after playback onset as shown by the echogram in Fig. 13.9b. Wild sprat (*Sprattus sprattus*) and mackerel (*Scomber scombrus*) exhibited a diversity of responses including break-up of aggregations and reforming of much denser aggregations in deeper water. The sprat is sensitive to sound pressure, however the mackerel lacks a swim bladder and is sensitive to the particle motion. The occurrence of behavioral responses increased with the received level. The 50% response thresholds were 163.2 and 163.3 dB re 1 µPa pk-pk and 135.0 and 142.0 dB re 1 $\mu Pa^2s$ (single-strike exposure) for sprat and mackerel, respectively (Hawkins et al. 2014; Fig. 13.10).

### 13.8.3 Effects of Noise on the Auditory and other Systems

After exposure to intense pulsed sound from airguns, extensive hearing damage in the form of ablated or missing hair cells was found in pink snapper (*Pagrus auratus*) (McCauley et al. 2003a, b). Other studies have found only limited or no hearing damage or threshold shift in various species of fish from airgun exposure (Hastings and Miksis-Olds 2012; Popper et al. 2005; Song et al. 2008). Apart from the typical differences in experimental setup, exposure regime, and species tested, a factor influencing the degree of noise impact might be the direction from which sound is received (specifically, vertical versus horizontal incidence; McCauley et al. 2003a). Fish ears are not symmetrical and many anthropogenic sound sources have a strong vertical directionality under water due to their near-surface deployment leading to a dipole sound field.

Halvorsen et al. (2012, Fig. 13.11) looked for tissue and organ damage in Chinook salmon (*Oncorhynchus tshawytscha*) that were placed inside a standing-wave test tube (High-Intensity

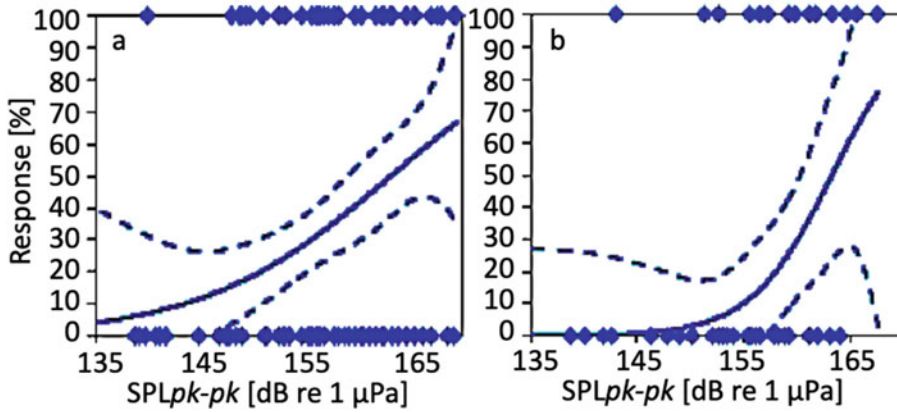



**Fig. 13.10** Dose-response curves (solid lines) and 95% confidence intervals (dashed lines) of (a) sprat and (b) mackerel to peak-to-peak sound pressure levels from pile driving (modified from Hawkins et al. 2014). © Acoustical Society of America, 2014. All rights reserved

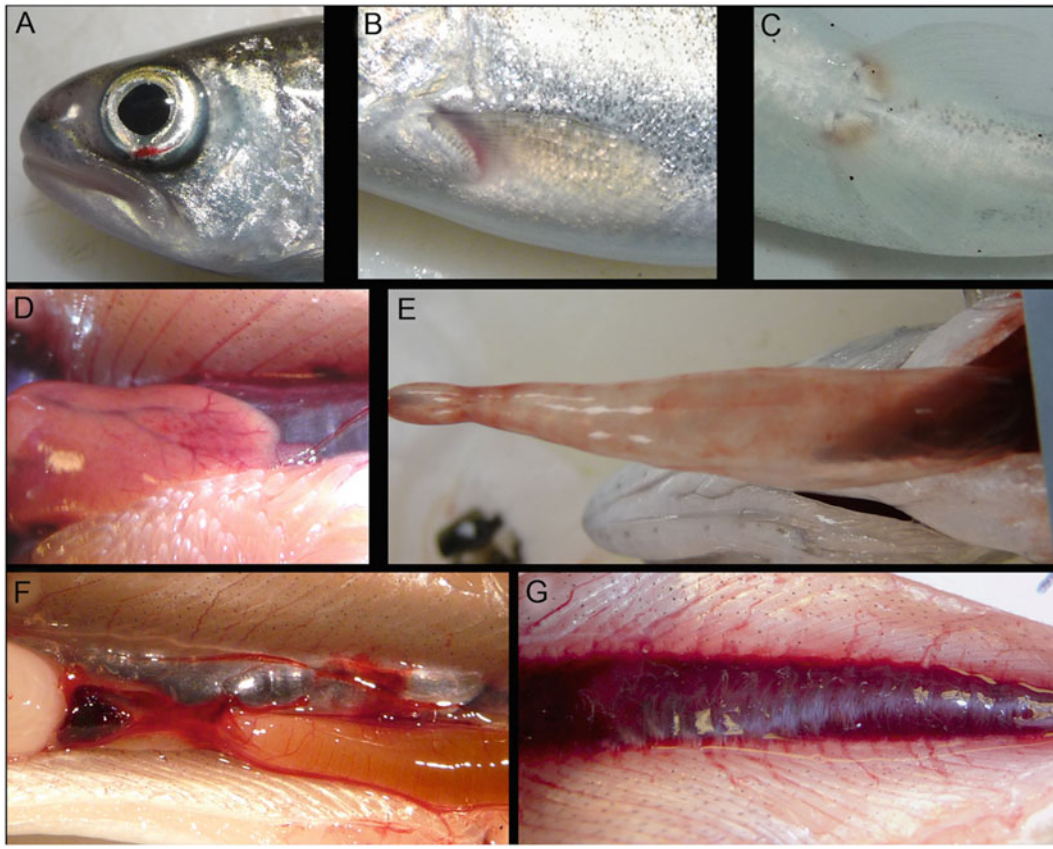



**Fig. 13.11** Chinook salmon injuries from noise exposure. Mild: (**a**) eye hemorrhage, (**b**, **c**) fin hematoma. Moderate: (**d**) liver hemorrhage and (**e**) bruised swim bladder. Mortal: (**f**) intestinal hemorrhage and (**g**) kidney hemorrhage (Halvorsen et al. 2012). © Halvorsen et al.; https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0038968; licensed under CC BY 4.0; https://creativecommons.org/licenses/by/4.0/

Controlled Impedance Fluid-filled wave Tube, HICI-FT) in which pressure and particle motion could be controlled. Physical injury commenced at 211 dB re 1 $\mu Pa^2s$ cumulative sound exposure resembling 1920 strikes of a pile driver at 177 dB re 1 $\mu Pa^2s$ each.

Yelverton (1975) conducted studies of the gross effects of sounds generated from underwater explosive blasts on fish. He found three important factors that influenced the degree of damage: the size of the fish relative to the wavelength of the sound, the species' anatomy, and the location of the fish in the water column relative to the sound source.

## 13.9 Noise Effects on Birds

Birds rely heavily on acoustic communication for life functions such as warning others about predators, finding and assessing the quality of mates, defending territories, and discerning which youngster to feed (Bradbury and Vehrencamp 2011). When environmental noise levels are high, such functions become difficult or impossible, unless the birds can make temporary or permanent adjustments to their signal, posture, or location. There have been several studies on the effects of noise on survival and communication in birds in the field as well as the laboratory, and on the ways that birds adjust their communication signals and/or lifestyles to adapt to the noisy modern world.

### 13.9.1 Bird Hearing

The avian ear has three main parts: an outer, middle, and inner ear. The outer ear is typically hidden by feathers, but consists of a small external meatus. A tympanic membrane separates the outer and middle ear. The middle ear contains the columella that mechanically transmits sound to the inner ear, and a connected interaural canal to aid in directional hearing. The basilar papilla in the inner ear converts sound into neural signals. Most birds hear between 50 Hz and 10 kHz, with some species' hearing extending into the infrasonic range (Dooling et al. 2000).

### 13.9.2 Behavioral Responses to Noise

Several studies have demonstrated that some birds are affected by low-frequency (<3 kHz) anthropogenic noise from roadways and that long-term exposure can lead to lower species diversity or lower breeding densities in an area (reviewed by Goodwin and Shriver 2011; Reijnen and Foppen 2006). Urban noise is known to affect reproduction and mating behaviors of birds in several ways. Urban noise can mask acoustic components of the lekking display by male greater sage grouse (*Centrocercus urophasianus*; Blickley and Patricelli 2012). It also disrupts female preference for low-frequency songs sung by male canaries (des Aunay et al. 2014) and great tits (Halfwerk et al. 2011). Females of these (and other) species prefer males that sing lower-frequency songs over those that sing higher-frequency songs because the low-frequency songs are sung by males of higher quality (e.g., Gil and Gahr 2002). When low-frequency urban noise masks the low-frequency components of calls and songs, females either cannot detect or find the males that are singing or cannot discriminate between the high-quality males singing at low frequencies and the poorer-quality males singing at higher frequencies.

Urban noise also has influences on where birds choose to live and breed, often resulting in consequences for choosing less favorable habitats. For instance, Eastern bluebirds (*Sialia sialis*) living in noisier environments were found to have reduced reproductive productivity and brood size compared to those living in quieter habitats (Kight et al. 2012). The presence and absence of construction and highways often changes the distribution of birds. Foppen and Deuzeman (2007) compared the distribution of reed warbler (*Acrocephalus arundinaceus*) pairs in the Netherlands before a highway was built through a nesting area and after the highway was present. When the highway was present

there were fewer nesting pairs, meaning that some birds were avoiding preferred habitats to avoid traffic noise. The road was temporarily closed and the number of nesting pairs increased; however, once the road reopened the number of nesting pairs again decreased. A more extensive study conducted in the Netherlands found that 26 of 43 (60%) woodland bird species showed reduced numbers near roads (Reijnen et al. 1995). Another count of birds near and far from roads showed that even when habitats were similar to one another, but either near to or far from a highway, the number of birds in each area increased with increasing distance from the road (Fig. 13.12), correlating with noise levels (Polak et al. 2013). That is, both abundance and diversity of birds increased as noise levels decreased. Other studies have confirmed that birds with higher-frequency calls were less likely to avoid the roadways than birds with lower-frequency calls (Rheindt 2003), again pointing to the challenges that many birds have when communicating in low-frequency urban noise, and highlighting the difficult choice that birds must face: Do the costs of choosing a less favorable habitat outweigh the benefits of living in quieter environments? The answer to this question clearly differs across both individual birds and species.

When birds do choose to nest in noisier environments, there could be consequences for mating and reproductive success. Nestling white-crowned sparrows (*Zonotrichia leucophrys*) tutored with songs embedded in anthropogenic noise later sung songs at higher frequencies and with lower vocal performance than those tutored with non-noisy control songs (Moseley et al. 2018). As another example, when alarm calls were presented to tree swallow (*Tachycineta bicolor*) nestlings, the tree swallows in quiet environments crouched more often (hiding from predators) while the nestlings in noisy environments produced longer calls and did not crouch (McIntyre et al. 2014). Nestling tree swallows living in noisier environments produced narrower-bandwidth and higher-frequency calls than those from quieter nests (Leonard and Horn 2008), although hearing of noise-reared nestlings does not differ from that of quiet-reared nestlings (Horn et al. 2020). These studies indicate that noise could affect how well offspring hear predators and how well parents hear begging calls. It also could influence the rate of feeding nestlings and could even have long-lasting effects on call structure, which could influence breeding success of those nestlings as adults. In a laboratory study looking at the effects of noise on reproduction, high levels of environmental noise eroded pair preferences in zebra finches (Swaddle and Page 2007). Paired females chose non-partner males over their partners when moderate to high

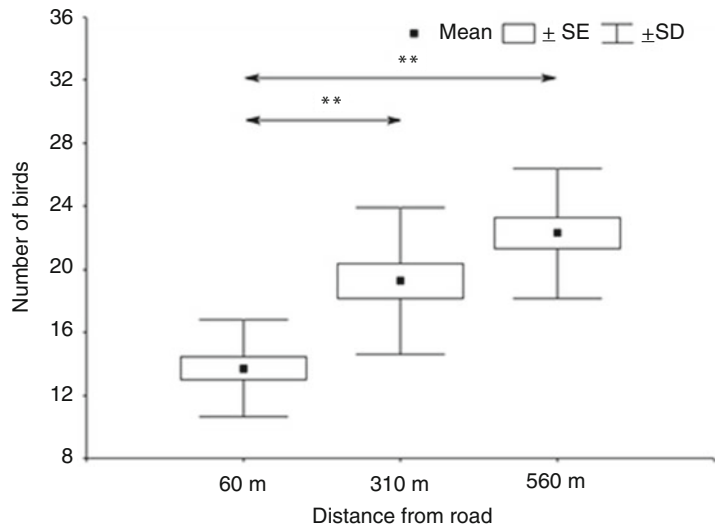



**Fig. 13.12** Relationship between bird abundance at point-count locations and distance from the road. Arrows show significant differences between points (Polak et al. 2013). © Springer Nature, 2013; https://link.springer.com/article/10.1007/s10342-013-0732-z/figures/5. Licensed under CC BY; https://creativecommons.org/licenses/

levels of white noise were presented in a preference test. These results have implications for noisy environments altering the population's breeding styles and eventually the evolutionary trajectory of the species (Swaddle and Page 2007).

### 13.9.3 Communication Masking

To know exactly how noise affects acoustic communication in birds, playback or perceptual experiments must be conducted to measure auditory acuity in a controlled environment. Experiments would use either pure tones and white noise or more complex and natural signals that birds use for communication purposes. Controlled laboratory studies measuring the ability to detect simple pure tones in broadband noise have been conducted in over a dozen bird species (reviewed by Dooling et al. 2000) using operant conditioning techniques. These studies have shown that as the frequency of the tone increases, it must be incrementally louder to hear it in a noisy background. This is not unlike the trend seen in other animals, suggesting a preserved evolutionary mechanism for hearing in noise.

Other laboratory studies measuring the detection and discrimination of calls and songs embedded in various types of noise can reveal more about the exact nature of the active space for the natural acoustic signals used for communication by social birds. Psychoacoustic studies often test the abilities of birds to detect, discriminate, or identify songs or calls that are embedded in a chorus of other songs or different types of noise (e.g., urban or woodland). Operant conditioning experiments on zebra finches, European starlings (*Sturnus vulgaris*), canaries (*Serinus canaria*), great tits (*Parus major*), and budgerigars all show that birds have excellent acuity for detecting or discriminating communication signals relative to pure tones, possibly due to the ecological relevance of these signals (Appeltants et al. 2005; Dent et al. 2009; Hulse et al. 1997; Lohr et al. 2003; Narayan et al. 2007; Pohl et al. 2009). In a field test of call discrimination, juvenile king penguins in a noisy colony were able to discriminate the calls of their parents from calls of other adults at a negative signal-to-noise ratio, suggesting that the enhanced detectability of natural vocal signals found in the laboratory actually translates to excellent acuity in the wild (Aubin and Jouventin 1998).

All of the above-mentioned studies reveal that songs and calls are more or less discriminable or detectable when they are presented within different masker types. For instance, great tits have better thresholds for detecting song elements embedded in woodland noise than urban noise (Fig. 13.13a; Pohl et al. 2009). Interestingly, detection of song elements in the dawn chorus was the most difficult condition for the great tits compared to the other noise types, suggesting that birds are not necessarily listening to one another in the mornings while they are singing. Canaries trained to identify canary songs embedded in one to four other distractor canary songs found it more difficult when there were more songs present, similar to conditions of the dawn chorus where many birds are singing overlapping songs (Fig. 13.13b; Appeltants et al. 2005). Another laboratory study determined birds' abilities to discriminate auditory distance, a task crucially important for territorial birds. Pohl et al. (2015) trained great tits to discriminate between virtual birdsongs at near and far distances, presented in quiet or embedded in a noisy dawn chorus. The birds accurately discriminated between distances, although this was much harder in noisy than in quiet conditions. In summary, these experiments and others demonstrate that hearing in noise is possible, and that factors such as the spectro-temporal make-up of signals, noise type, and noise level all have an influence on hearing signals in noise.

As a whole, results from the laboratory and field experiments suggest that bird communication is more successful in quiet, rather than noisy environments, that the type of noise matters for communication, and that if noise is present, adjustments need to be made to the calls or songs of signalers for those signals to be detected, discriminated, and localized by the receivers. One such adjustment that has shown to be effective is changing the position of the signal relative to the

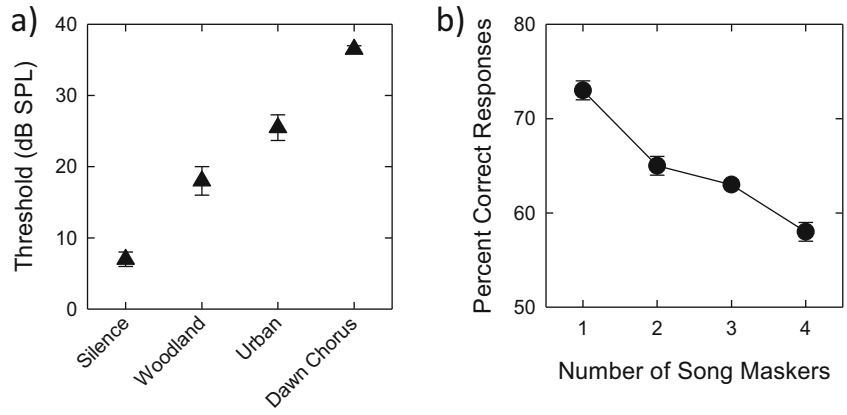



**Fig. 13.13** (**a**) Masked thresholds for great tits detecting a synthetic song element embedded in silence, woodland noise, urban noise, or dawn chorus noise (adapted from Pohl et al. 2009). Performance is best for quiet conditions, worst for the chorus conditions. Thresholds are higher for urban noise than woodland noise. (**b**) Performance for canaries discriminating song elements embedded in 1–4 other songs (adapted from Appeltants et al. 2005). As the number of maskers increases, performance decreases

masker. Dent et al. (1997) found that thresholds for budgerigars detecting a pure tone in white noise were 11 dB lower when the signal and noise were separated by 90° in space than when they were co-located (i.e., spatial release from masking). A follow-up study showed an even greater advantage when the spatially separated signal was zebra finch song and the masker was a zebra finch chorus (Fig. 13.14; Dent et al. 2009). Thus, when birds are trying to communicate with one another in noisy environments, changing their position or even simply moving their heads will increase communication efficiency in similar ways as humans attempting to speak to one another in a noisy cocktail party will often move their head toward a speaker.

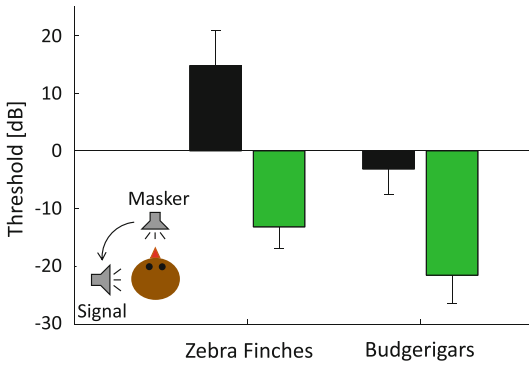



**Fig. 13.14** Signal-to-noise ratio thresholds for detecting a zebra finch song are higher (worse) when a chorus masker is co-located with the song (black boxes) than when the song is spatially separated from the masker (green boxes), in both budgerigars and zebra finches. Adapted from Dent et al. (2009)

Another adjustment made by many birds is to shift the frequency content of songs to a higher range, as documented for European blackbirds (*Turdus merula*; Slabbekoorn and Ripmeester 2008), plumbeous vireos (*Vireo plumbeus*; Francis et al. 2011), gray vireos (*Vireo vicinior*; Francis et al. 2011), European robins (McMullen et al. 2014), chaffinches (Verzijden et al. 2010), black-capped chickadees (*Poecile atricapillus*; Proppe et al. 2011), and a number of tropical birds (de Magalhães Tolentino et al. 2018). Whether this is a true adaptation attempting to increase the lowest frequencies of songs above the highest frequencies of the noise, whether it is simply easier for the birds to make high frequencies louder, or whether urban birds live in denser environments and want to distinguish their songs from those of other birds is still being debated (e.g., Nemeth et al. 2013).

Pohl et al. (2012) tested the consequences of such shifts on perception in the laboratory. These authors trained great tits to detect or discriminate

between song phrases embedded in urban or woodland noises. In the urban noise background, it was easier for the tits to detect the high-frequency phrases than the low-frequency phrases. There was no difference in the woodland noise for detection of the different song types. For birds attempting to discriminate high- or low-frequency songs embedded in woodland or urban noises, the researchers found that the high-frequency elements were more useful in urban conditions, while the whole song was used for discrimination in woodland noise. Thus, birds that are changing their calls and songs into higher-frequency ranges for improved communication in noisy urban environments are doing so adaptively.

Other vocal adjustments made by birds in response to noise are to sing more during the quiet night than during the noisy day (as in European robins; Fuller et al. 2007), to shift the initiation of the dawn chorus by as much as 5 h to compensate for traffic noise (as in European blackbirds; Nordt and Klenke 2013), and to increase the intensity of vocalizations (Lombard effect). Black-capped chickadees modify the structure and frequencies of their alarm calls in response to noise (Courter et al. 2020), while house wrens (*Troglodytes aedon*) reduce the size of their song repertoires in addition to changing their song frequencies (Juárez et al. 2021). In a field study on noisy miners (*Manorina melanocephala*), Lowry et al. (2012) found that individuals at noisier locations produced louder alarm calls than those at quieter locations. The Lombard effect has also been demonstrated in the laboratory in Japanese quail (*Coturnix japonica*; Potash 1972), budgerigars (Manabe et al. 1998), chickens (*Gallus gallus domesticus*; Brumm et al. 2009), nightingales (*Luscinia megarhynchos*; Brumm and Todt 2002), white-rumped munia (*Lonchura striata*; Kobayasi and Okanoya 2003), and zebra finches (Cynx et al. 1998). A recent experiment measuring songs of the white-crowned sparrows in urban San Francisco during the 2020 COVID-19 shutdown showed that the birds responded to the decrease in noise levels with a return to decades-old song frequencies (Derryberry et al. 2020), suggesting that they have an almost-immediate ability to re-occupy an acoustic niche within a soundscape.

### 13.9.4 Physiological Effects

One major advantage birds possess, compared to humans, is the ability to regenerate auditory sensory cells lost during exposure to very loud sounds (Ryals and Rubel 1988), therefore birds experience no hearing loss over time from either aging or noisy environments. Birds do, however, experience stress from noise (Blickley et al. 2012; Strasser and Heath 2013).

Acoustic communication in birds is vital for survival, and understanding how noise affects sound production and perception is important for conservation efforts. Birds are clearly affected by the increasing levels of urban noise in their environments, but many adjust their calling and singing styles or locations to overcome problems of communicating in noise. Certainly, there are both limits to and consequences of those adjustments.

## 13.10 Noise Effects on Terrestrial Mammals

Anthropogenic noise affects mammals in a variety of ways changing their behavior, physiology, and ultimately ability to succeed in what otherwise might be considered optimal habitat. Terrestrial mammals show responses that range from ignoring or tolerating to avoiding noise, with potential impacts ranging from negligible to severe (Slabbekoorn et al. 2018b).

### 13.10.1 Terrestrial Mammal Hearing

Among terrestrial mammals, humans (*Homo sapiens*) are the most studied species with prevalent research addressing hearing physiology and psychology, hearing loss, and restoration. The mammalian ear consists of mechanical structures (incus, malleus, and stapes) evolutionarily derived from elements of the jaw that function

to translate sound from acoustic waves to nerve signals in the cochlea and auditory nerve. Though very effective, the ear can sustain damage and it degrades with age. Hearing loss results in reduced auditory acuity and limited information for the mammal to use. Loss can be caused by sudden exposure to high-intensity sound (e.g., from an explosion or gunfire) or by repeated or prolonged noise exposure (e.g., at industrial workplaces, at rock concerts, or from personal media players).

While the general structure of the mammalian ear is shared amongst terrestrial mammal species, there is great diversity in the sounds mammals can perceive, in the sounds they produce, and in their responses to sound. While human hearing ranges from about 20 Hz to 20 kHz, elephants use infrasound (sounds extending below the human hearing range, i.e., below 20 Hz; Herbst et al. 2012; Payne et al. 1986) and bats use ultrasound (sounds extending above the human hearing range, i.e., above 20 kHz, with some species hearing and emitting sound up to 220 kHz; Fenton et al. 2016). Rodents are known to be quite diverse, with subterranean species having excellent low-frequency hearing and terrestrial rodents having excellent ultrasonic hearing (reviewed by Dent et al. 2018). Mammals can thus be expected to display a diversity of responses to noise.

### 13.10.2 Behavioral Responses to Noise

One of the most frequently studied sources of noise in terrestrial mammal habitats is traffic noise from cars, trains, or aircraft. The most frequently reported response is animal movement away from the noise source. For example, Sonoran pronghorn (*Antilocapra americana sonoriensis*) increased their use of areas with lower levels of noise over areas with higher levels of noise from military aircraft (Landon et al. 2003). In the case of mountain sheep (*Ovis canadensis mexicana*), 19% showed disturbance to low-flying aircraft (Krausman and Hervert 1983). Prairie dogs (*Cynomys ludovicianus*) were exposed to playback of highway noise in an experimental prairie-dog town that was previously absent of anthropogenic noise. The treatment area had fewer prairie dogs above ground. Those that were above ground spent less time foraging and much more time exhibiting vigilant behavior (Shannon et al. 2014) leading to earlier predator detection and earlier flight response (Shannon et al. 2016).

A major concern regarding these behavioral responses by wildlife to traffic corridors is habitat fragmentation together with limited connectivity. Noisy areas may displace wildlife and form barriers to migration and dispersal (Barber et al. 2011; Fig. 13.15). Roads also fragment bat

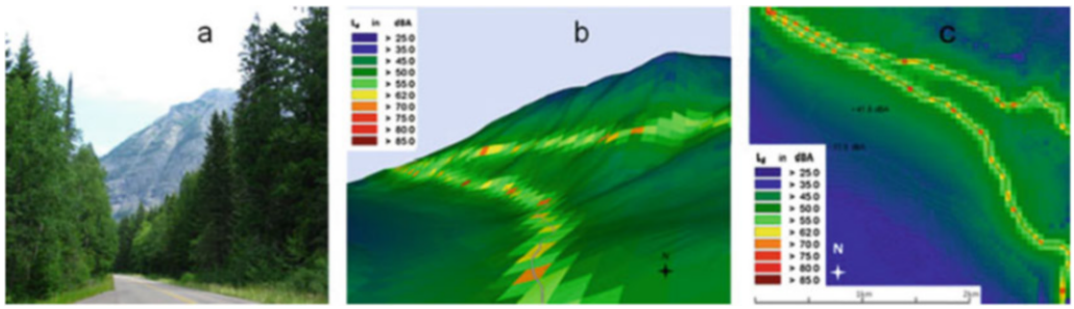



**Fig. 13.15** (**a**) Photo of the Going-to-the-Sun road in Glacier National Park, USA. (**b**) 3D plot of 24-h traffic noise. (**c**) 2D plot of 24-h traffic noise (Barber et al. 2011). Road noise may form a barrier to wildlife migration. Reprinted by permission from Springer Nature. Barber, J. R., Burdett, C. L., Reed, S. E., Warner, K. A., Formichella, C., Crooks, K. R., Theobald, D. M., and Fristrup, K. M. Anthropogenic noise exposure in protected natural areas: estimating the scale of ecological consequences. Landscape Ecology, 26(9), 1281; https://link.springer.com/article/10.1007/s10980-011-9646-7. © Springer Nature, 2011. All rights reserved

habitat, although many species cross roadways or fly through underpasses (Kerth and Melber 2009).

Animals may adapt temporal behavioral patterns around noise exposure. Black-tufted marmosets (*Callithrix penicillata*) living in an urban park in Brazil stayed in quieter, central (i.e., away from road noise) areas during the day, and only utilized the park edges at night or weekends (Duarte et al. 2011). Forest elephants (*Loxodonta cyclotis*) became more nocturnal in areas of industrial activity; and while the study found no direct link to noise intensity, concern about natural biorhythms near noisy industrial sites was raised (Wrege et al. 2010).

Noise may affect foraging behavior. Woodland caribou stopped feeding when exposed to noise from petroleum exploration (Bradshaw et al. 1997). Reduced food intake in noise slowed growth in rats, pigs, and dogs (Alario et al. 1987; Gue et al. 1987; Otten et al. 2004). Gleaning bats (*Myotis myotis*) displayed reduced hunting efficiency during road noise playbacks (Schaub et al. 2008; Siemers and Schaub 2011). Similarly, Brazilian free-tailed bats (*Tadarida brasiliensis*) were less active and produced fewer echolocation bursts near a noisy gas compression station (Bunkley et al. 2015). Peromyscus mice, on the other hand, were more successful collecting pine seeds (a major food source) near noisy gas-extraction sites because competing, seed-collecting jays (*Aphelocoma californica*) abandoned the site (Francis et al. 2012). Additionally, predators of the mice, like owls, avoided the noisier sites, which may result in reduced predation of the mice (Mason et al. 2016). Finally, some animals may associate noise with reinforcement, such as food sources, and learn to approach sounds. Badgers (*Meles meles*) quickly learned to approach an acoustic deterrent device baited with food (dinner bell effect; Ward et al. 2008).

One pathway by which noise disrupts animal behavior is by acoustic masking. Piglets use vocalization bouts to coordinate nursing with sows and noise disrupted this communication leading to reduced milk ingestion and increased energetic costs for the piglets attempting to elicit milk (Algers and Jensen 1985). Some animals can adjust their calls to reduce masking (Lombard effect). Cats increased the amplitude of calls in noise (Nonaka et al. 1997). Common marmosets (*Callithrix jacchus*) and cotton-top tamarins (*Saguinus oedipus*) increased both amplitude and duration of calls in noise (Brumm et al. 2004; Roian Egnor and Hauser 2006). Cotton-top tamarins timed their calls to avoid overlap with periodic noise (Egnor et al. 2007). Horseshoe bats (*Rhinolophidae*) increased echolocation amplitudes and shifted echolocation frequency in noise (Hage et al. 2013).

### 13.10.3 Physiological Responses to Noise

Human studies have shown that noise exposure can lead to a variety of health effects ranging from a feeling of annoyance to disturbed sleep, emotional stress, decreased job performance, higher chance of developing cardiovascular disease, and decreased learning in schoolchildren (Basner et al. 2014). We can only begin to understand the effects of noise on the health of other mammalian species.

Studies on elk (*Cervus canadensis*) and wolves (*Canis lupus*) in Yellowstone National Park, USA, had elevated levels of glucocorticoid enzymes (a blood hormone that indicates stress) when snowmobiles were allowed in the park. After banning snowmobiling, enzyme levels returned to normal, although a direct link to noise exposure was not made (Creel et al. 2002). After ongoing zoo visitor noise, giant pandas (*Ailuropoda melanoleuca*) exhibited increased glucocorticoids, negatively impacting reproduction efforts (Owen et al. 2004). In male rats exposed to chronic noise, testosterone decreased (Ruffoli et al. 2006). Pregnant mice exposed to 85–95 dB re 20 μPa alarm bells had pups with lower serum IgG levels, indicating impaired immune responses (Sobrian et al. 1997). Chronic noise exposure in rats affected calcium regulation leading to detrimental changes at cellular level (Gesi et al. 2002). Desert mule deer (*Odocoileus hemionus crooki*) and mountain sheep had increased heart rates relative to increased levels of aircraft noise playback. Heart rate returned to

normal within 60–180 s and responses decreased over time potentially indicating a form of habituation (Weisenberger et al. 1996).

### 13.10.4 Effects of Noise on the Auditory System

The physiological impact of noise is well documented in several mammalian species, particularly laboratory animals, due to the ability to systematically expose and test individuals. Systematic research has shown that several sound features (such as sound frequency, duration, intensity, amplitude rise time, continuous versus temporary exposure, etc.) impact how an animal's auditory system is affected by noise exposure. For example, chinchillas experienced TTS from exposure to the sound of a hammer hitting a nail repeatedly (Dunn et al. 1991). While some of the chinchillas were exposed to repeated hammering (a series of separate sound events), others were exposed to continuous noise of the same spectrum as nail hammering (one single sound event). While all chinchillas showed a decrease in hearing sensitivity, the chinchillas exposed to the repeated hammering had more hearing loss (Dunn et al. 1991).

NIHL can occur from mechanical damage and/or from metabolic disruption of acoustic structures (Hu 2012). Mechanical damage occurs during the sound exposure due to excessive movement caused by sound waves. Depending on the level of the sound, loud noise can damage structures at the cellular level. Metabolic damage occurs due to a cascade of changes at the cellular level from mechanical damage and can continue for weeks after sound exposure.

In TTS, damage may occur to the synapses and stereocilia, while in PTS, damage is more extensive, including outer hair cell death and fibrocyte loss. For example, the audiograms of four species of Old-World monkeys (*Macaca nemestrina, M. mulatta, M. fascicularis*, and *Papio papio*) were compared before and after exposure to octave-band noise (between 0.5 and 8 kHz at levels of 120 dB re 20 μPa) for 8 h daily for 20 days. Loss of both inner and outer hair cells at the basal end of the organ of Corti and hence PTS were produced (Hawkins et al. 1976). The difference in noise exposure when an individual transitions from having temporary to permanent damage varies by species as well as depending on several individual factors such as past sound exposure, age, genetics, etc. (Hu 2012).

Exposure to continuous, high-level (>100 dB re 20 μPa) sounds has been shown to damage or destroy hair cells in multiple species, such as rats, rabbits, and guinea pigs (Borg et al. 1995; Chen and Fechter 2003; Hu et al. 2000). Recently, exposure to lower-amplitude sounds over long periods of time has also been shown to cause permanent damage. Mice exposed to 70 dB re 20 μPa continuous white noise for 8 h a day over the course of up to 3 months showed increased hearing thresholds and decreased auditory response amplitudes (Feng et al. 2020). Notably, the mice also showed aggravated age-related hearing loss in relatively young mice (mice were 8 weeks old at the start of exposure) (Feng et al. 2020).

Some animals can mitigate the impact of noise on the auditory system using a stapedial reflex to close the auditory meatus. When exposed to a loud sound, the contraction of the stapedial muscle causes a decrease in auditory sensitivity by closing the auditory meatus, thus negating some potential damage. This reflex is well documented in humans and appears to primarily play a role in sudden, unexpected sounds with sharp rise times. The reflex is thought to function similarly in most terrestrial mammals, for example in rabbits. Rabbits exposed to sound in normal conditions had very little threshold shifts, but when their stapedial reflex was inactivated (by blocking the nerve) during noise exposure, PTS was observed at otherwise not NIHL inducing levels (Borg et al. 1983). In cats, this reflex functions even under anesthesia (McCue and Guinan 1994). However, damage to the auditory nerve connections (synaptopathy) can also damage auditory reflexes; for example, in mice, synaptopathy was directly correlated to the function of the middle ear muscle reflex (Valero et al. 2018).

Synaptopathy not only occurs from noise exposure, but also at old age or from exposure to ototoxins (Valero et al. 2018).

## 13.11 Noise Effects on Marine Mammals

As with terrestrial animals, the potential effects of noise exposure on marine mammals may include a range of physical effects on auditory and other systems, as well as behavioral responses, and interference with sound communication systems (Erbe et al. 2018; Southall 2018). Several reviews have recently been completed, for specific noise sources (such as shipping, Erbe et al. 2019b; dredging, Todd et al. 2015; and wind farms, Madsen et al. 2006), and specific geographic regions (such as Antarctica; Erbe et al. 2019a). Current knowledge is summarized here, ranging from issues that are likely most experienced, but less severe, to effects that may more rarely occur but are increasingly severe. Events of the latter category, such as mass strandings and mortalities of marine mammals associated with strong acute anthropogenic sounds (notably certain military active sonar systems or explosives), have historically driven and dominated the awareness, interest, and research on the potential effects of noise on marine mammals (e.g., Filadelfo et al. 2009). However, there is increasing concern over sub-lethal, yet potentially more widespread, effects (notably behavioral influences) of more chronic noise sources and their consequences for individual fitness and ultimately population parameters (e.g., New et al. 2014; Ocean Studies Board 2016). Southall et al. (2007) reviewed the available literature at that time and made specific recommendations regarding effects of anthropogenic noise on hearing and behavior in marine mammals. Substantial additional research and synthesis of available data has expanded on their assessment, improving the empirical basis for these evaluations and expanding consideration to other important areas discussed here (e.g., masking and auditory impact thresholds; Erbe et al. 2016a; Finneran 2015). And so the Southall et al. (2007) criteria were updated in 2019 (Southall et al. 2019b).

### 13.11.1 Marine Mammal Hearing

In most situations of noise exposure, marine mammals might merely detect a sound without a specific adverse effect. Furthermore, animals arguably have to be able to detect signals in order for most of the effects described here to potentially occur. Hearing capabilities and specializations vary widely in marine mammals. Some species, such as pinnipeds, have adaptations to facilitate both aerial and underwater hearing (Reichmuth et al. 2013). Other species, including the odontocete cetaceans, have very wide frequency ranges of underwater hearing extending well into ultrasonic ranges to facilitate echolocation (Mooney et al. 2012). For other key species, including many of the endangered mysticete cetaceans, virtually no direct data are available regarding hearing, which is instead estimated from anatomical and sound production parameters.

Southall et al. (2007) developed the concept of functional marine mammal hearing groups. Each group was assigned a frequency-specific auditory filter (called weighting function) to account for known and presumed differences in hearing sensitivity within marine mammals (Fig. 13.16). Using additional direct data, these hearing groups and weighting functions were substantially improved and modified (Finneran 2016). These weighting functions are applied to the noise spectrum in order to estimate the likelihood of NIHL, by comparison to published TTS and PTS onset thresholds expressed as weighted cumulative sound exposure levels (National Marine Fisheries Service 2018).

Understanding and directly accounting for the frequency-specific parameters of noise and how they interact with background noise and marine mammal-specific hearing is important in considering the contextual aspects of potential behavioral responses (Ellison et al. 2012), auditory masking (Erbe et al. 2016a), and hearing impairment and damage (e.g., Finneran 2015).

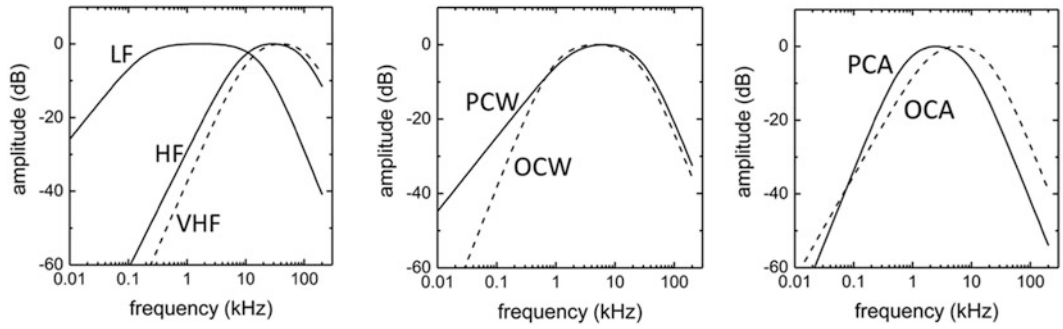



**Fig. 13.16** Auditory weighting functions for marine mammal functional hearing groups; LF: low-frequency cetaceans, HF: high-frequency cetaceans, VHF: very-high-frequency cetaceans, PCW: phocid carnivores in water, OCW: other carnivores in water, PCA: phocid carnivores in air, OCA: other carnivores in air (Southall et al. 2019b)

### 13.11.2 Behavioral Responses to Noise

Noise exposure may lead to a variety of behavioral responses (and severity) in marine mammals, ranging from minor changes in orientation to separation of mothers and dependent offspring, or mass mortality. Southall et al. (2007) reviewed these responses and proposed a qualitative relative severity scaling that takes into account the relative duration and potential impacts on biologically meaningful activities. This approach has been applied and modified in quantifying behavioral responses in the context of exposure-response risk functions (e.g., Miller et al. 2012; Southall et al. 2019a). While sound exposure level is an important aspect of determining the relative probability of a response, other contextual factors of exposure also may be critically important, including animal behavioral state (e.g., Goldbogen et al. 2013), spatial proximity to the noise (e.g., Ellison et al. 2012), sensitization to noise exposure (Kastelein et al. 2011), or nearby vessel noise (Dunlop et al. 2020). A variety of experimental and observational methods have been applied in evaluating noise exposure and behavioral responses, resulting in a large volume of scientific literature on this subject that is reviewed generally here.

Behavioral responses to noise have been studied in both field and laboratory. The advantage of field studies is the observation of animals in their natural environment, but it can be challenging to observe individuals and determine exposure levels and responses with sufficient resolution and sample size. Field studies of large sample size include observations of changes in whale distribution in response to industrial noise and seismic surveys (see Richardson et al. 1995 for an overview), recordings of vocal behavior of whales exposed to military sonar (Fristrup et al. 2003; Miller et al. 2000), and a recent series of experiments exposing migrating humpback whales to 20, 440, and 3300-in$^3$ seismic airgun arrays (Dunlop et al. 2016, 2017a, 2020). Many recent experimental field studies have considered potential effects of active sonar on cetaceans (Southall et al. 2016). Among the many broad results and conclusions are dose-response curves for exposure level and response probability in killer whales (Miller et al. 2014) and humpback whales (Dunlop et al. 2017b, 2018), behavioral state-dependent responses in blue whales (*Balaenoptera musculus*; Goldbogen et al. 2013) and humpback whales (Dunlop et al. 2017a, 2020), and changes in social behavior following noise exposure in pilot whales (*Globicephala* sp.; Visser et al. 2016) and humpback whales (Dunlop et al. 2020). For instance, Goldbogen et al. (2013) showed that deep-feeding blue whales are much more likely to change diving behavior and body orientation in response to noise than those in shallow-feeding or non-feeding states

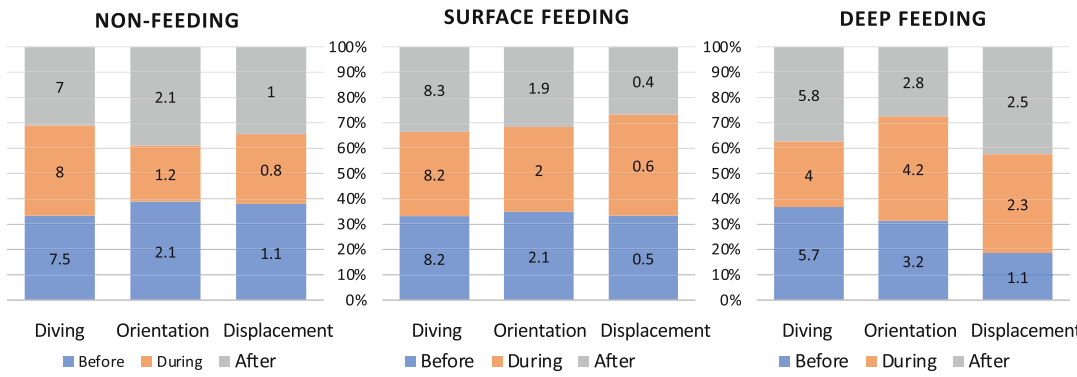



**Fig. 13.17** Relative response differences in various aspects of blue whale behavior between non-feeding, surface-feeding, and deep-feeding individuals (adapted from Goldbogen et al. 2013). Response magnitude was quantified using generalized additive mixed models for behavioral parameters relevant to each behavioral state and potential responses in terms of diving, orientation, and displacement

(Fig. 13.17). This finding has been replicated and expanded with individual blue whales, demonstrating the same context-dependency in response probability as well as potential dependence in response probability based on horizontal range from the sound source even for the same received levels (Southall et al. 2019a).

Some species such as long-finned pilot whales appear behaviorally tolerant of noise exposure (e.g., Antunes et al. 2014), whereas beaked whales (Family Ziphiidae) are clearly among the more sensitive species behaviorally (DeRuiter et al. 2013; Miller et al. 2015; Stimpert et al. 2014; Tyack et al. 2011). The analysis of multivariate behavioral data to determine changes in behavior, including potentially subtle but important changes, is statistically challenging, although recent substantial progress in analytical methods has been made as well (Harris et al. 2016).

Experimental laboratory approaches have the advantage of greater control and precision on multivariate aspects of exposure and response, but lack the contextual reality in which free-ranging animals experience noise. Studies that evaluated noise exposure and response probability in captive harbor porpoises (e.g., Kastelein et al. 2011, 2013) demonstrated a particular sensitivity of this species, which matched field observations. Studies with captive bottlenose dolphins (*Tursiops truncatus*) and California sea lions (*Zalophus californianus*) have included large sample sizes and repeated exposures to demonstrate species, age, and experiential differences in response probability to military sonar signals (Houser et al. 2013a, b).

Observational methods (visual and acoustic) have provided complementary data to assess both acute and chronic noise exposure. Passive acoustic monitoring over large areas and time periods demonstrated changes in acoustic behavior and inferred movement of beaked whales in response to military sonar signals (e.g., McCarthy et al. 2011) resulting in dose-response curves (Moretti et al. 2014). Similarly, large-scale monitoring linked cetacean distribution and behavior to seismic surveys (e.g., Pirotta et al. 2014; Thompson et al. 2013), impact pile driving (e.g., Dähne et al. 2013; Thompson et al. 2010; Tougaard et al. 2009), and acoustic harassment devices (e.g., Johnston 2002).

Such observational studies lack experimental control, resolution to the individual level, detail on fine-scale responses, and ability to differentiate short-term responses to noise from those to other stimuli, but offer information on broad-scale spatio-temporal changes in habitat use and behavior. Ideally, experimental approaches would be combined with broad-scale observational methods to discover potential population-level effects (see Southall et al. 2016).

### 13.11.3 Communication Masking

Noise can interfere with or "mask" acoustic communication by marine mammals (Erbe et al. 2016a). Masking is due to the simultaneous presence of signal and noise energy within the same frequency bands. Masking reduces the range over which a signal may be detected. Or, in other words, the signal must be louder, for it to be detected in the presence of noise (Fig. 13.18).

The area over which an animal call can be detected by its intended recipients (i.e., the active space or communication space) fluctuates in space and time. Models have been developed to quantify lost communication space and applied to mysticetes communicating near busy shipping lanes (Fig. 13.19; Clark et al. 2009; Hatch et al. 2012).

The Lombard effect has been demonstrated in marine mammals as an increase in vocalization source levels (e.g., Helble et al. 2020; Holt et al. 2009; Thode et al. 2020), duration (Miller et al. 2000), or repetition (Thode et al. 2020). Additionally, marine mammals have demonstrated increased detection capabilities based on angular separation between signal and noise sources, termed a spatial release from masking (e.g., Turnbull 1994), or based on wide-band amplitude-modulation patterns in the noise, termed a comodulation masking release (e.g., Branstetter et al. 2013). These compensatory and signal processing capabilities reduce the masking potential of noise.

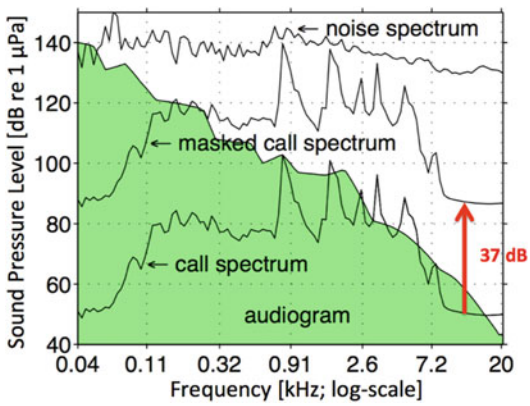



**Fig. 13.18** Beluga whale (*Delphinapterus leucas*) audiogram (shaded green), spectrum of a call at detection threshold (measured behaviorally) in the absence of noise, spectrum of an icebreaker's bubbler noise, and the masked call spectrum in the presence of bubbler noise. The spectra are shown as band levels, with the bandwidths aiming to represent the auditory filters. The upwards shift of the call spectrum equals the amount of masking: 37 dB (Erbe 2000)

### 13.11.4 Effects of Noise on the Auditory and Other Systems

While behavioral responses and auditory masking may occur relatively far from sound sources, impacts to the auditory system are expected at higher levels hence shorter ranges. As with masking, the frequency of noise exposure is important in terms of the potential for NIHL, and noise at frequencies where animals are more sensitive has a greater potential for inducing such effects in marine mammals (Finneran 2015). Furthermore, the temporal pattern of noise matters substantially in terms of the potential for NIHL. Impulsive signals with rapid rise times are more likely to cause NIHL (see Finneran 2015). The risk and severity of NIHL increases with repeated and longer exposures, but simple energy-based models integrating exposure level over time cannot fully predict potential NIHL.

Despite substantial recent research, our understanding of NIHL in marine mammals remains limited. TTS has been studied in fewer than ten species, and not in any mysticete. Controlled exposure experiments that would produce a PTS are infeasible due to animal ethics considerations. Nonetheless, TTS studies in odontocetes and pinnipeds produced TTS-onset levels and information on frequency-dependence (reviewed by Finneran 2015). Recent experiments produced frequency-weighted TTS-onset levels higher than the original exposure criteria compiled by Southall et al. (2007). However, some studies (e.g., Kastelein et al. 2012; Lucke et al. 2009) demonstrated much lower TTS-onset levels, specifically in harbor porpoises.

Noise may further cause non-auditory physiological impacts that may not be immediately apparent. Noise has increased stress hormones in

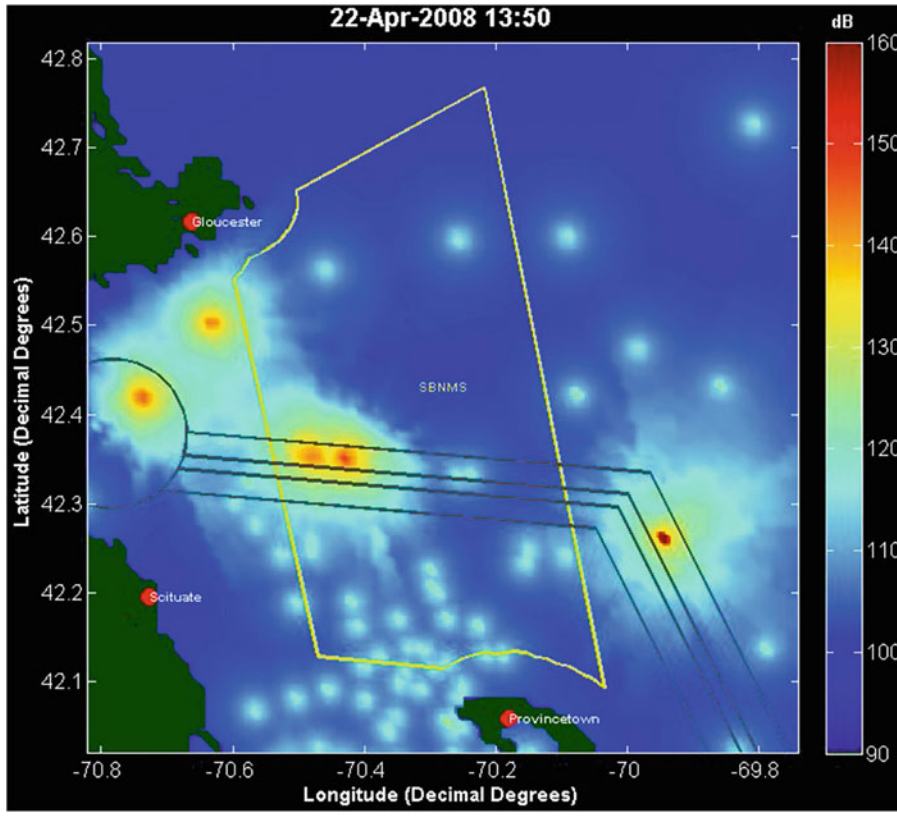



**Fig. 13.19** Chart of acoustic footprints of North Atlantic right whales (*Eubalaena glacialis*; light blue dots) and ships (larger footprints with red centers) off Cape Cod, Massachusetts Bay, USA. The larger and stronger ship noise footprints can easily engulf (i.e., mask) the right whale calls. Stellwagen Bank National Marine Sanctuary outlined in yellow. Figure courtesy of Chris Clark

the blood of captive marine mammals (e.g., Romano et al. 2004). In the wild, stress hormones in right whales decreased when ambient noise from shipping was lower (Rolland et al. 2012). Such measurements of noise-induced stress in marine mammals are comparable to studies with other vertebrates (Romero and Butler 2007). However, information is lacking on how stress scales with noise exposure and on the long-term health impacts of prolonged stress.

Finally, beaked whales that stranded after exposure to military sonar exhibited lesions and gas or fat emboli (Fernandez et al. 2005; Jepson et al. 2003). While some form of decompression sickness has been hypothesized, the physiological mechanisms for such emboli to occur are poorly understood. These physiological effects may have been secondarily caused or exacerbated by the animals' behavioral responses to sonar.

## 13.12 Summary

This chapter presented examples of the variety of effects noise can have on animals in terrestrial and aquatic habitats. Studies on the hearing in noise and on behavioral and physiological responses to noise have concentrated on fish, frogs, birds, terrestrial mammals, and marine mammals. Clearly, more research is needed for invertebrates, reptiles, and all groups of freshwater species. In addition, more studies on the metabolic costs of these responses are needed.

Animals demonstrate a hierarchy of behavioral and physiological responses to noise. Behavioral reactions to anthropogenic noise include a startle response, change in movement and direction, freezing in place, cessation of vocal behavior, and change in behavioral budgets. Animals can also modify their signals to counteract the effects of noise and improve communication. Such modifications include changes in amplitude, duration, and frequency. Some animals also increase the redundancy of their signals by repeating them more often. Physiological reactions to anthropogenic noise are indicated by increased cortisol levels (indication of stress), temporary or permanent hearing loss, and physical damage to tissues and organs such as lungs and swim bladders.

The effects of anthropogenic noise on individual animals can escalate to the population level. Ultimately, species-richness and biodiversity could be affected. However, methods and models to address these topics are in their infancy.

There is the potential to mitigate any negative impacts of anthropogenic noise by modifying the noise source characteristics and operation schedules, finding alternative means to obtain operational goals of the noise source, and protecting critical habitats. Effective management of habitats should include noise assessment. Further research is needed to understand the ecological consequences of chronic noise in terrestrial and aquatic environments.

Remote wilderness areas are not immune to the effects of anthropogenic noise, because sound travels very well (with little loss over long ranges) in many terrestrial and aquatic habitats. Resource managers should continue to be vigilant in monitoring and mitigating the effects of anthropogenic noise on animals.

## References


Alario P, Gamallo A, Beato MJ, Trancho G (1987) Body weight gain, food intake and adrenal development in chronic noise stressed rats. Physiol Behav 40(1): 29–32. https://doi.org/10.1016/0031-9384(87)90181-8

Algers B, Jensen P (1985) Communication during suckling in the domestic pig. Effects of continuous noise. Appl Anim Behav Sci 14(1):49–61. https://doi.org/10.1016/0168-1591(85)90037-1

American National Standards Institute (2013) Acoustical Terminology (ANSI/ASA S1.1-2013). Acoustical Society of America, Melville, NY

Amézquita A, Hödl W (2004) How, when, and where to perform visual displays: The case of the Amazonian frog *Hyla parviceps*. Herpetologica 60(4):420–429. https://doi.org/10.1655/02-51

André M, Solé M, Lenoir M, Durfort M, Quero C, Mas A, Lombarte A, Mvd S, López-Bejar M, Morell M, Zaugg S, Houégnigan L (2011) Low-frequency sounds induce acoustic trauma in cephalopods. Front Ecol Environ 9(9):489–493. https://doi.org/10.1890/100124

Andriguetto-Filho JM, Ostrensky A, Pie MR, Silva UA, Boeger WA (2005) Evaluating the impact of seismic prospecting on artisanal shrimp fisheries. Cont Shelf Res 25(14):1720–1727. https://doi.org/10.1016/j.csr.2005.05.003

Antunes R, Kvadsheim PH, Lam FPA, Tyack PL, Thomas L, Wensveen PJ, Miller PJO (2014) High thresholds for avoidance of sonar by free-ranging long-finned pilot whales (*Globicephala melas*). Mar Pollut Bull 83(1):165–180. https://doi.org/10.1016/j.marpolbul.2014.03.056

Appeltants D, Gentner TQ, Hulse SH, Balthazart J, Ball GF (2005) The effect of auditory distractors on song discrimination in male canaries (*Serinus canaria*). Behav Process 69(3):331–341. https://doi.org/10.1016/j.beproc.2005.01.010

Arkhipkin AI, Bizikov VA (2000) Role of the statolith in functioning of the acceleration receptor system in squids and sepioids. J Zool 250(1):31–55

Au WWL, Floyd RW, Penner RH, Murchison AE (1974) Measurement of echolocation signals of the Atlantic bottlenose dolphin, *Tursiops truncatus* Montagu, in open waters. J Acoust Soc Am 56(4):1280–1290

Au WWL, Carder DA, Penner RH, Scronce BL (1985) Demonstration of adaptation in beluga whale echolocation signals. J Acoust Soc Am 77(2):726–730. https://doi.org/10.1121/1.392341

Aubin T, Jouventin P (1998) Cocktail–party effect in king penguin colonies. Proc R Soc Lond Ser B Biol Sci 265(1406):1665

Bain D, Dahlheim M (1994) Effects of masking noise on detection thresholds of killer whales. In: Loughlin T (ed) Marine mammals and the Exxon Valdez. Academic Press, San Diego, CA, pp 243–256

Barber JR, Crooks KR, Fristrup KM (2010) The costs of chronic noise exposure for terrestrial organisms. Trends Ecol Evol 25(3):180–189. https://doi.org/10.1016/j.tree.2009.08.002

Barber JR, Burdett CL, Reed SE, Warner KA, Formichella C, Crooks KR, Theobald DM, Fristrup KM (2011) Anthropogenic noise exposure in protected natural areas: estimating the scale of ecological consequences. Landsc Ecol 26(9):1281. https://doi.org/10.1007/s10980-011-9646-7

Barnett KE, Cocroft RB, Fleishman LJ (1999) Possible communication by substrate vibration in a chameleon. Copeia 1:225–228. https://doi.org/10.2307/1447408

Basner M, Babisch W, Davis A, Brink M, Clark C, Janssen S, Stansfeld S (2014) Auditory and non-auditory effects of noise on health. Lancet 383(9925):1325–1332. https://doi.org/10.1016/S0140-6736(13)61613-X

Battershill C, Cappo M, Colquhoun J, Cripps E, Jorgensen D, McCorry D, Stowar M, Venables W (2008) Coral damage monitoring using Towed Video (TVA) and Photo Quadrat Assessments (PQA). Australian Institute of Marine Science, Townsville, QLD

Bee MA (2007) Sound source segregation in grey treefrogs: spatial release from masking by the sound of a chorus. Anim Behav 74(3):549–558. https://doi.org/10.1016/j.anbehav.2006.12.012

Bee MA, Swanson EM (2007) Auditory masking of anuran advertisement calls by road traffic noise. Anim Behav 74(6):1765–1776. https://doi.org/10.1016/j.anbehav.2007.03.019

Bergen F, Abs M (1997) Etho-ecological study of the singing activity of the Blue Tit (*Parus caeruleus*), Great Tit (*Parus major*) and Chaffinch (*Fringilla coelebs*). J Ornithol 138(4):451–467. https://doi.org/10.1007/bf01651380

Blackwell SB, Nations CS, TL MD, Thode AM, Mathias D, Kim KH, Green CR, Macrander AM (2015) Effects of airgun sounds on bowhead whale calling rates: evidence for two behavioral thresholds. PLoS One 10(6):e0125720. https://doi.org/10.1371/journal.pone.0125720

Blickley JL, Patricelli GL (2012) Potential acoustic masking of greater sage-grouse (*Centrocercus urophasianus*) display components by chronic industrial noise. Ornithol Monogr 74:23–35. https://doi.org/10.1525/om.2012.74.1.23

Blickley JL, Word KR, Krakauer AH, Phillips JL, Sells SN, Taff CC, Wingfield JC, Patricelli GL (2012) Experimental chronic noise is related to elevated fecal corticosteroid metabolites in lekking male greater Sage-Grouse (*Centrocercus urophasianus*). PLoS One 7(11):e50462. https://doi.org/10.1371/journal.pone.0050462

Bohne T, Grießmann T, Rolfes R (2019) Modeling the noise mitigation of a bubble curtain. J Acoust Soc Am 146(4):2212–2223. https://doi.org/10.1121/1.5126698

Bolm-Audorff U, Hegewald J, Pretzsch A, Freiberg A, Nienhaus A, Seidler A (2020) Occupational noise and hypertension risk: a systematic review and meta-analysis. Int J Environ Res Public Health 17(17):6281. https://doi.org/10.3390/ijerph17176281

Borg E, Nilsson R, Engström B (1983) Effect of the acoustic reflex on inner ear damage induced by industrial noise. Acta Otolaryngol 96(5-6):361–369. https://doi.org/10.3109/00016488309132721

Borg E, Canlon B, Engström B (1995) Noise-induced hearing loss. Literature review and experiments in rabbits. Morphological and electrophysiological features, exposure parameters and temporal factors, variability and interactions. Scand Audiol Suppl 40:1–147

Bowles AE, Eckert S, Starke L, Berg E, Wolski L (1999) Effects of flight noise from jet aircraft and sonic booms on hearing, behavior, heart rate and oxygen consumption of desert tortoises (*Gopherus agassizii*). Hubbs-Sea World Research Institution, San Diego, CA

Bradbury JW, Vehrencamp SL (2011) Principles of animal communication, 2nd edn. Sinauer Associates, Sunderland, MA

Bradshaw CJA, Boutin S, Hebert DM (1997) Effects of petroleum exploration on Woodland Caribou in Northeastern Alberta. J Wildl Manag 61(4):1127–1133. https://doi.org/10.2307/3802110

Branstetter BK, Trickey JS, Bakhtiari K, Black A, Aihara H, Finneran JJ (2013) Auditory masking patterns in bottlenose dolphins (*Tursiops truncatus*) with natural, anthropogenic, and synthesized noise. J Acoust Soc Am 133(3):1811–1818. https://doi.org/10.1121/1.4789939

Brittan-Powell EF, Christensen-Dalsgaard J, Tang Y, Carr C, Dooling RJ (2010) The auditory brainstem response in two lizard species. J Acoust Soc Am 128(2):787–794. https://doi.org/10.1121/1.3458813

Brumm H (ed) (2013) Animal communication and noise. Animal signals and communication, vol 2. Springer, Berlin. https://doi.org/10.1007/978-3-642-41494-7

Brumm H, Slabbekoorn H (2005) Acoustic communication in noise. In: Advances in the study of behavior, vol 35. Academic Press, New York, pp 151–209. https://doi.org/10.1016/S0065-3454(05)35004-2

Brumm H, Todt D (2002) Noise-dependent song amplitude regulation in a territorial songbird. Anim Behav 63(5):891–897. https://doi.org/10.1006/anbe.2001.1968

Brumm H, Voss K, Köllmer I, Todt D (2004) Acoustic communication in noise: regulation of call characteristics in a New World monkey. J Exp Biol 207(3):443. https://doi.org/10.1242/jeb.00768

Brumm H, Schmidt R, Schrader L (2009) Noise-dependent vocal plasticity in domestic fowl. Anim Behav 78(3):741–746. https://doi.org/10.1016/j.anbehav.2009.07.004

Bunkley JP, McClure CJW, Kleist NJ, Francis CD, Barber JR (2015) Anthropogenic noise alters bat activity levels and echolocation calls. Global Ecol Conserv 3:62–71. https://doi.org/10.1016/j.gecco.2014.11.002

Chan AAY-H, Giraldo-Perez P, Smith S, Blumstein DT (2010) Anthropogenic noise affects risk assessment and attention: the distracted prey hypothesis. Biol Lett 6(4):458–461. https://doi.org/10.1098/rsbl.2009.1081

Chen G-D, Fechter LD (2003) The relationship between noise-induced hearing loss and hair cell loss in rats. Hear Res 177(1):81–90. https://doi.org/10.1016/S0378-5955(02)00802-X

Christensen CB, Christensen-Dalsgaard J, Brandt C, Madsen PT (2012) Hearing with an atympanic ear: good vibration and poor sound-pressure detection in the royal python, *Python regius*. J Exp Biol 215(2): 331. https://doi.org/10.1242/jeb.062539

Clark CW, Ellison WT, Southall BL, Hatch L, Van Parijs SM, Frankel A, Ponirakis D (2009) Acoustic masking in marine ecosystems: intuitions, analysis, and implication. Mar Ecol Prog Ser 395:201–222. https://doi.org/10.3354/Meps08402

Cocroft RB, Rodríguez RL (2005) The behavioral ecology of insect vibrational communication. Bioscience 55(4): 323–334. https://doi.org/10.1641/0006-3568(2005)055[0323:TBEOIV]2.0.CO;2

Cook SL, Forrest TG (2005) Sounds produced by nesting leatherback sea turtles (*Dermochelys coriacea*). Herpetol Rev 36(4):387–390

Coram A, Gordon J, Thompson D, Northridge SP (2014) Evaluating and Assessing the relative effectiveness of acoustic deterrent devices and other non-lethal measures on marine mammals. University of St Andrews, Sea Mammal Research Unit, St Andrews, Scotland

Costa DP, Schwarz L, Robinson P, Schick RS, Morris PA, Condit R, Crocker DE, Kilpatrick AM (2016) A bioenergetics approach to understanding the population consequences of disturbance: elephant seals as a model system. In: Popper AN, Hawkins T (eds) The effects of noise on aquatic life II. Springer, New York, pp 161–169. https://doi.org/10.1007/978-1-4939-2981-8_19

Costello RA, Symes LB (2014) Effects of anthropogenic noise on male signalling behaviour and female phonotaxis in Oecanthus tree crickets. Anim Behav 95:15–22. https://doi.org/10.1016/j.anbehav.2014.05.009

Courter JR, Perruci RJ, McGinnis KJ, Rainieri JK (2020) Black-capped chickadees (*Poecile atricapillus*) alter alarm call duration and peak frequency in response to traffic noise. PLoS One 15(10):e0241035. https://doi.org/10.1371/journal.pone.0241035

Creel S, Fox JE, Hardy A, Sands J, Garrott B, Peterson RO (2002) Snowmobile activity and glucocorticoid stress responses in wolves and elk. Conserv Biol 16(3): 809–814. https://doi.org/10.1046/j.1523-1739.2002.00554.x

Cunnington GM, Fahrig L (2010) Plasticity in the vocalizations of anurans in response to traffic noise. Acta Oecol 36(5):463–470. https://doi.org/10.1016/j.actao.2010.06.002

Cynx J, Lewis R, Tavel B, Tse H (1998) Amplitude regulation of vocalizations in noise by a songbird, *Taeniopygia guttata*. Anim Behav 56(1):107–113. https://doi.org/10.1006/anbe.1998.0746

Dähne M, Gilles A, Lucke K, Peschko V, Adler S, Krügel K, Sundermeyer J, Siebert U (2013) Effects of pile-driving on harbour porpoises (*Phocoena phocoena*) at the first offshore wind farm in Germany. Environ Res Lett 8(2):025002. https://doi.org/10.1088/1748-9326/8/2/025002

Davidsen JG, Dong H, Linné M, Andersson MH, Piper A, Prystay TS, Hvam EB, Thorstad EB, Whoriskey F, Cooke SJ, Sjursen AD, Rønning L, Netland TC, Hawkins AD (2019) Effects of sound exposure from a seismic airgun on heart rate, acceleration and depth use in free-swimming Atlantic cod and saithe. Conserv Physiol 7(1). https://doi.org/10.1093/conphys/coz020

Davis AK, Schroeder H, Yeager I, Pearce J (2018) Effects of simulated highway noise on heart rates of larval monarch butterflies, *Danaus plexippus*: implications for roadside habitat suitability. Biol Lett 14(5): 20180018. https://doi.org/10.1098/rsbl.2018.0018

Day RD, McCauley RD, Fitzgibbon QP, Hartmann K, Semmens JM (2016a) Assessing the impact of marine seismic surveys on southeast Australian scallop and lobster fisheries. Fisheries Research & Development Corporation

Day RD, McCauley RD, Fitzgibbon QP, Semmens JM (2016b) Seismic air gun exposure during early-stage embryonic development does not negatively affect spiny lobster *Jasus edwardsii* larvae (Decapoda: Palinuridae). Sci Rep 6:22723. https://doi.org/10.1038/srep22723

Day RD, McCauley RD, Fitzgibbon QP, Hartmann K, Semmens JM (2017) Exposure to seismic air gun signals causes physiological harm and alters behavior in the scallop *Pecten fumatus*. Proc Natl Acad Sci 114(40):E8537. https://doi.org/10.1073/pnas.1700564114

Day RD, McCauley RD, Fitzgibbon QP, Hartmann K, Semmens JM (2019) Seismic air guns damage rock lobster mechanosensory organs and impair righting reflex. Proc Biol Sci 286(1907):20191424. https://doi.org/10.1098/rspb.2019.1424

Day RD, Fitzgibbon QP, McCauley RD, Hartmann K, Semmens JM (2020) Lobsters with pre-existing damage to their mechanosensory statocyst organs do not incur further damage from exposure to seismic air gun signals. Environ Pollut. https://doi.org/10.1016/j.envpol.2020.115478

de Magalhães Tolentino VC, Baesse CQ, Melo C (2018) Dominant frequency of songs in tropical bird species is higher in sites with high noise pollution. Environ Pollut 235:983–992. https://doi.org/10.1016/j.envpol.2018.01.045

de Soto NA, Delorme N, Atkins J, Howard S, Williams J, Johnson M (2013) Anthropogenic noise causes body malformations and delays development in marine larvae. Sci Rep 3:2831. https://doi.org/10.1038/srep02831

Dent ML, Larsen ON, Dooling RJ (1997) Free-field binaural unmasking in budgerigars (*Melopsittacus undulatus*). Behav Neurosci 111(3):590–598. https://doi.org/10.1037/0735-7044.111.3.590

Dent ML, McClaine EM, Best V, Ozmeral E, Narayan R, Gallun FJ, Sen K, Shinn-Cunningham BG (2009) Spatial unmasking of birdsong in zebra finches (*Taeniopygia guttata*) and budgerigars (*Melopsittacus undulatus*). J Comp Psychol 123(4):357–367. https://doi.org/10.1037/a0016898

Dent ML, Screven LA, Kobrina A (2018) Hearing in Rodents. In: Dent ML, Fay RR, Popper AN (eds) Rodent bioacoustics. Springer, Cham, pp 71–105. https://doi.org/10.1007/978-3-319-92495-3_4

Department of the Navy (2008) Atlantic fleet active sonar training environmental impact statement. Naval Facilities Engineering Command, Atlantic, Norfolk, VA

Derryberry EP, Phillips JN, Derryberry GE, Blum MJ, Luther D (2020) Singing in a silent spring: Birds respond to a half-century soundscape reversion during the COVID-19 shutdown. Science 370(6516):575. https://doi.org/10.1126/science.abd5777

DeRuiter SL, Larbi Doukara K (2012) Loggerhead turtles dive in response to airgun sound exposure. Endanger Species Res 16(1):55–63. https://doi.org/10.3354/esr00396

DeRuiter SL, Southall BL, Calambokidis J, Zimmer WMX, Sadykova D, Falcone EA, Friedlaender AS, Joseph JE, Moretti D, Schorr GS, Thomas L, Tyack PL (2013) First direct measurements of behavioural responses by Cuvier's beaked whales to mid-frequency active sonar. Biol Lett 9(4). https://doi.org/10.1098/rsbl.2013.0223

des Aunay GH, Slabbekoorn H, Nagle L, Passas F, Nicolas P, Draganoiu TI (2014) Urban noise undermines female sexual preferences for low-frequency songs in domestic canaries. Anim Behav 87:67–75. https://doi.org/10.1016/j.anbehav.2013.10.010

Doksaeter L, Godo OR, Handegard NO (2009) Behavioural responses of herring (*Clupea harengus*) to 1-2 and 6-7 kHz sonar signals and killer whale feedings sounds. J Acoust Soc Am 125(1):554–564. https://doi.org/10.1121/1.3021301

Dooling RJ, Leek MR (2018) Communication masking by man-made noise. In: Slabbekoorn H, Dooling RJ, Popper AN, Fay RR (eds) Effects of anthropogenic noise on animals. Springer, New York, pp 23–46. https://doi.org/10.1007/978-1-4939-8574-6_2

Dooling RJ, Lohr B, Dent ML (2000) Hearing in birds and reptiles. In: Dooling RJ, Fay RR, Popper AN (eds) Comparative hearing: birds and reptiles. Springer, New York, pp 308–359. https://doi.org/10.1007/978-1-4612-1182-2_7

Duarte MHL, Vecci MA, Hirsch A, Young RJ (2011) Noisy human neighbours affect where urban monkeys live. Biol Lett 7(6):840–842. https://doi.org/10.1098/rsbl.2011.0529

Dunlop RA, Cato DH, Noad MJ (2014) Evidence of a Lombard response in migrating humpback whales (*Megaptera novaeangliae*). J Acoust Soc Am 136(1):430–437. https://doi.org/10.1121/1.4883598

Dunlop RA, Noad MJ, McCauley RD, Kniest E, Slade R, Paton D, Cato DH (2016) Response of humpback whales (*Megaptera novaeangliae*) to ramp-up of a small experimental air gun array. Mar Pollut Bull 103(1–2):72–83. https://doi.org/10.1016/j.marpolbul.2015.12.044

Dunlop RA, Noad MJ, McCauley RD, Kniest E, Slade R, Paton D, Cato DH (2017a, 1869) The behavioural response of migrating humpback whales to a full seismic airgun array. Proc Biol Sci 284. https://doi.org/10.1098/rspb.2017.1901

Dunlop RA, Noad MJ, McCauley RD, Scott-Hayward L, Kniest E, Slade R, Paton D, Cato DH (2017b) Determining the behavioural dose–response relationship of marine mammals to air gun noise and source proximity. J Exp Biol 220(16):2878–2886. https://doi.org/10.1242/jeb.160192

Dunlop RA, Noad MJ, McCauley RD, Kniest E, Slade R, Paton D, Cato DH (2018) A behavioural dose-response model for migrating humpback whales and seismic air gun noise. Mar Pollut Bull 133:506–516. https://doi.org/10.1016/j.marpolbul.2018.06.009

Dunlop RA, McCauley RD, Noad MJ (2020) Ships and air guns reduce social interactions in humpback whales at greater ranges than other behavioral impacts. Mar Pollut Bull 154:111072. https://doi.org/10.1016/j.marpolbul.2020.111072

Dunn DE, Davis RR, Merry CJ, Franks JR (1991) Hearing loss in the chinchilla from impact and continuous noise exposure. J Acoust Soc Am 90(4):1979–1985. https://doi.org/10.1121/1.401677

Egnor SER, Wickelgren JG, Hauser MD (2007) Tracking silence: adjusting vocal production to avoid acoustic interference. J Comp Physiol A 193(4):477–483. https://doi.org/10.1007/s00359-006-0205-7

Ellison W, Southall B, Clark C, Frankel A (2012) A new context-based approach to assess marine mammal behavioral responses to anthropogenic sounds. Conserv Biol 26(1):21–28. https://doi.org/10.1111/j.1523-1739.2011.01803.x

Engas A, Løkkeborg S (2002) Effects of seismic shooting and vessel-generated noise, on fish behaviour and catch rates. Bioacoustics 12(2–3):313–316

Engås A, Løkkeborg S, Ona E, Soldal AV (1996) Effects of seismic shooting on local abundance and catch rates of cod (*Gadus morhua*) and haddock (*Melanogrammus aeglefinus*). Can J Fish Aquat Sci 53(10):2238–2249. https://doi.org/10.1139/f96-177

Erbe C (2000) Detection of whale calls in noise: performance comparison between a beluga whale, human listeners and a neural network. J Acoust Soc Am 108(1):297–303. https://doi.org/10.1121/1.429465

Erbe C, Reichmuth C, Cunningham KC, Lucke K, Dooling RJ (2016a) Communication masking in marine mammals: a review and research strategy. Mar Pollut Bull 103:15–38. https://doi.org/10.1016/j.marpolbul.2015.12.007

Erbe C, Sisneros J, Thomsen F, Hawkins A, Popper A (2016b) Overview of the fourth international conference on the effects of noise on aquatic life. Proc Meet Acoust 27(1):010006. https://doi.org/10.1121/2.0000256

Erbe C, Dunlop R, Dolman S (2018) Effects of noise on marine mammals. In: Slabbekoorn H, Dooling RJ, Popper AN, Fay RR (eds) Effects of anthropogenic

noise on animals. Springer, New York, pp 277–309. https://doi.org/10.1007/978-1-4939-8574-6_10

Erbe C, Dähne M, Gordon J, Herata H, Houser DS, Koschinski S, Leaper R, McCauley R, Miller B, Müller M, Murray A, Oswald JN, Scholik-Schlomer AR, Schuster M, van Opzeeland IC, Janik VM (2019a) Managing the effects of noise from ship traffic, seismic surveying and construction on marine mammals in Antarctica. Front Mar Sci 6:647. https://doi.org/10.3389/fmars.2019.00647

Erbe C, Marley S, Schoeman R, Smith JN, Trigg L, Embling CB (2019b) The effects of ship noise on marine mammals: a review. Front Mar Sci 6:606. https://doi.org/10.3389/fmars.2019.00606

Erbe C, Sisneros J, Thomsen F, Lepper P, Hawkins A, Popper A (2019c) Overview of the fifth international conference on the effects of noise on aquatic life. Proc Meet Acoust 37(1):001001. https://doi.org/10.1121/2.0001052

Fay RR, Popper AN (2012) Fish hearing: new perspectives from two 'senior' bioacousticians. Brain Behav Evol 79(4):215–217. https://doi.org/10.1159/000338719

Feng AS, Narins PM (2008) Ultrasonic communication in concave-eared torrent frogs (*Amolops tormotus*). J Comp Physiol A 194(2):159–167. https://doi.org/10.1007/s00359-007-0267-1

Feng S, Yang L, Hui L, Luo Y, Du Z, Xiong W, Liu K, Jiang X (2020) Long-term exposure to low-intensity environmental noise aggravates age-related hearing loss via disruption of cochlear ribbon synapses. Am J Transl Res 12(7):3674–3687

Fenton MB, Grinnell AD, Popper AN, Fay RR (eds) (2016) Bat bioacoustics. Springer handbook of auditory research, vol 54. Springer, New York

Fernandez A, Edwards JF, Rodriguez F, de los Monteros AE, Herraez P, Castro P, Jaber JR, Martin V, Arbelo M (2005) "Gas and fat embolic syndrome" involving a mass stranding of beaked whales (Family Ziphiidae) exposed to anthropogenic sonar signals. Vet Pathol 42(4):446–457

Ferrara CR, Vogt RC, Harfush MR, Sousa-Lima RS, Albavera E, Tavera A (2014) First evidence of leatherback turtle (*Dermochelys coriacea*) embryos and hatchlings emitting sounds. Chelonian Conserv Biol 13(1):110–114. https://doi.org/10.2744/ccb-1045.1

Fewtrell JL, McCauley RD (2012) Impact of air gun noise on the behaviour of marine fish and squid. Mar Pollut Bull 64(5):984–993. https://doi.org/10.1016/j.marpolbul.2012.02.009

Fields DM, Handegard NO, Dalen J, Eichner C, Malde K, Karlsen Ø, Skiftesvik AB, Durif CMF, Browman HI (2019) Airgun blasts used in marine seismic surveys have limited effects on mortality, and no sublethal effects on behaviour or gene expression, in the copepod *Calanus finmarchicus*. ICES J Mar Sci 76(7):2033–2044. https://doi.org/10.1093/icesjms/fsz126

Filadelfo R, Mintz J, Michlovich E, Amico AD, Tyack PL, Ketten DR (2009) Correlating military sonar use with beaked whale mass strandings: what do the historical data show? Aquat Mamm 35:435–444. https://doi.org/10.1578/AM.35.4.2009.435

Finneran JJ (2015) Noise-induced hearing loss in marine mammals: a review of temporary threshold shift studies from 1996 to 2015. J Acoust Soc Am 138(3):1702–1726. https://doi.org/10.1121/1.4927418

Finneran JJ (2016) Auditory weighting functions and TTS/PTS exposure functions for marine mammals exposed to underwater noise. National Marine Fisheries Service, Silver Spring, MD

Fitzgibbon QP, Day RD, McCauley RD, Simon CJ, Semmens JM (2017) The impact of seismic air gun exposure on the haemolymph physiology and nutritional condition of spiny lobster, *Jasus edwardsii*. Mar Pollut Bull 125(1):146–156. https://doi.org/10.1016/j.marpolbul.2017.08.004

Foppen RPB, Deuzeman S (2007) De Grote karekiet in de noordelijke randmeren; een dilemma voor natuurontwikkelingsplannen!? De Levende Natuur 108:20–26

Francis CD, Ortega CP, Cruz A (2009) Noise pollution changes avian communities and species interactions. Curr Biol 19(16):1415–1419. https://doi.org/10.1016/j.cub.2009.06.052

Francis CD, Ortega CP, Cruz A (2011) Different behavioural responses to anthropogenic noise by two closely related passerine birds. Biol Lett 7(6):850–852. https://doi.org/10.1098/rsbl.2011.0359

Francis CD, Kleist NJ, Ortega CP, Cruz A (2012) Noise pollution alters ecological services: enhanced pollination and disrupted seed dispersal. Proc R Soc B Biol Sci 279(1739):2727–2735. https://doi.org/10.1098/rspb.2012.0230

Fristrup KM, Hatch LT, Clark CW (2003) Variation in humpback whale (*Megaptera novaeangliae*) song length in relation to low-frequency sound broadcasts. J Acoust Soc Am 113(6):3411–3424. https://doi.org/10.1121/1.1573637

Fuller RA, Warren PH, Gaston KJ (2007) Daytime noise predicts nocturnal singing in urban robins. Biol Lett 3(4):368. https://doi.org/10.1098/rsbl.2007.0134

Galeotti P, Sacchi R, Fasola M, Ballasina D (2005) Do mounting vocalisations in tortoises have a communication function? A comparative analysis. Herpetol J 15(2):61–71

Gallego-Abenza M, Mathevon N, Wheatcroft D (2019) Experience modulates an insect's response to anthropogenic noise. Behav Ecol 31(1):90–96. https://doi.org/10.1093/beheco/arz159

Gesi M, Fornai F, Lenzi P, Ferrucci M, Soldani P, Ruffoli R, Paparelli A (2002) Morphological alterations induced by loud noise in the myocardium: the role of benzodiazepine receptors. Microsc Res Tech 59(2):136–146. https://doi.org/10.1002/jemt.10186

Gil D, Gahr M (2002) The honesty of bird song: multiple constraints for multiple traits. Trends Ecol Evol 17(3): 133–141. https://doi.org/10.1016/S0169-5347(02)02410-2

Goldbogen JA, Southall BL, DeRuiter SL, Calambokidis J, Friedlaender AS, Hazen EL, Falcone EA, Schorr GS, Douglas A, Moretti DJ, Kyburg C, McKenna MF, Tyack PL (2013) Blue whales respond to simulated mid-frequency military sonar. Proc R Soc B 280(1765):20130657. https://doi.org/10.1098/rspb.2013.0657

Goodwin SE, Shriver WG (2011) Effects of traffic noise on occupancy patterns of forest birds. Conserv Biol 25(2):406–411. https://doi.org/10.1111/j.1523-1739.2010.01602.x

Grafe TU, Preininger D, Sztatecsny M, Kasah R, Dehling JM, Proksch S, Hödl W (2012) Multimodal communication in a noisy environment: A case study of the Bornean rock frog *Staurois parvus*. PLoS One 7(5): e37965. https://doi.org/10.1371/journal.pone.0037965

Greenfield MD (2016) Evolution of acoustic communication in insects. In: Pollack GS, Mason AC, Popper AN, Fay RR (eds) Insect hearing. Springer, Cham, pp 17–47. https://doi.org/10.1007/978-3-319-28890-1_2

Gue M, Fioramonti J, Frexinos J, Alvinerie M, Bueno L (1987) Influence of acoustic stress by noise on gastrointestinal motility in dogs. Dig Dis Sci 32(12): 1411–1417. https://doi.org/10.1007/BF01296668

Guerra A, González AF, Rocha F (2004) A review of the records of giant squid in the north-eastern Atlantic and severe injuries in *Architeuthis dux* stranded after acoustic explorations. In: Paper presented at the ICES Annual Science Conference, Vigo, Spain

Hage SR, Jiang T, Berquist SW, Feng J, Metzner W (2013) Ambient noise induces independent shifts in call frequency and amplitude within the Lombard effect in echolocating bats. Proc Natl Acad Sci 110(10):4063. https://doi.org/10.1073/pnas.1211533110

Hahad O, Kröller-Schön S, Daiber A, Münzel T (2019) The cardiovascular effects of noise. Dtsch Arztebl Int 116(14):245–250. https://doi.org/10.3238/arztebl.2019.0245

Halfwerk W, Bot S, Buikx J, van der Velde M, Komdeur J, ten Cate C, Slabbekoorn H (2011) Low-frequency songs lose their potency in noisy urban conditions. Proc Natl Acad Sci 108(35):14549–14554. https://doi.org/10.1073/pnas.1109091108

Halfwerk W, Lea AM, Guerra MA, Page RA, Ryan MJ (2016) Vocal responses to noise reveal the presence of the Lombard effect in a frog. Behav Ecol 27(2): 669–676. https://doi.org/10.1093/beheco/arv204

Halvorsen MB, Casper BM, Woodley CM, Thomas J, Carlson TJ, Popper AN (2012) Threshold for onset of injury in Chinook salmon from exposure to impulsive pile driving sounds. PLoS One 7(6):e38968. https://doi.org/10.1371/journal.pone.0038968

Harris CM, Thomas L, Sadykova D, DeRuiter SL, Tyack PL, Southall BL, Read AJ, Miller PJO (2016) The challenges of analyzing behavioral response study data: an overview of the MOCHA (Multi-study OCean Acoustics Human Effects Analysis) project. In: Popper AN, Hawkins A (eds) The effects of noise on aquatic life II. Springer, New York, pp 399–407. https://doi.org/10.1007/978-1-4939-2981-8_47

Hastings MC, Miksis-Olds J (2012) Shipboard assessment of hearing sensitivity of tropical fishes immediately after exposure to seismic air gun emissions at Scott Reef. In: Popper AN, Hawkins A (eds) The effects of noise on aquatic life. Springer, New York, pp 239–243. https://doi.org/10.1007/978-1-4419-7311-5_53

Hatch L, Clark C, Van Parijs S, Frankel A, Ponirakis D (2012) Quantifying loss of acoustic communication space for right whales in and around a U.S. National Marine Sanctuary. Conserv Biol 26(6):983–994. https://doi.org/10.1111/j.1523-1739.2012.01908.x

Hawkins AD, Popper AN (2018) Effects of man-made sound on fishes. In: Slabbekoorn H, Dooling RJ, Popper AN, Fay RR (eds) Effects of anthropogenic noise on animals. Springer, New York, pp 145–177. https://doi.org/10.1007/978-1-4939-8574-6_6

Hawkins JE, Johnsson LG, Stebbins WC, Moody DB, Coombs SL (1976) Hearing loss and cochlear pathology in monkeys after noise exposure. Acta Otolaryngol 81(3-6):337–343. https://doi.org/10.3109/00016487609119971

Hawkins AD, Roberts L, Cheesman S (2014) Responses of free-living coastal pelagic fish to impulsive sounds. J Acoust Soc Am 135(5):3101–3116. https://doi.org/10.1121/1.4870697

Helble TA, Guazzo RA, Martin CR, Durbach IN, Alongi GC, Martin SW, Boyle JK, Henderson EE (2020) Lombard effect: Minke whale boing call source levels vary with natural variations in ocean noise. J Acoust Soc Am 147(2):698–712. https://doi.org/10.1121/10.0000596

Herbst CT, Stoeger AS, Frey R, Lohscheller J, Titze IR, Gumpenberger M, Fitch WT (2012) How low can you go? Physical production mechanism of elephant infrasonic vocalizations. Science 337(6094):595–599. https://doi.org/10.1126/science.1219712

Herrera-Montes MI, Aide TM (2011) Impacts of traffic noise on anuran and bird communities. Urban Ecosyst 14(3):415–427. https://doi.org/10.1007/s11252-011-0158-7

Heuch PA, Karlsen E (1997) Detection of infrasonic water oscillations by copepodids of *Lepeophtheirus salmonis* (Copepoda Caligida). J Plankton Res 19(6):735–747. https://doi.org/10.1093/plankt/19.6.735

Heyward A, Colquhoun J, Cripps E, McCorry D, Stowar M, Radford B, Miller K, Miller I, Battershill C (2018) No evidence of damage to the soft tissue or skeletal integrity of mesophotic corals exposed to a 3D marine seismic survey. Mar Pollut Bull 129(1):8–13. https://doi.org/10.1016/j.marpolbul.2018.01.057

Hofstetter RW, Dunn DD, McGuire R, Potter KA (2014) Using acoustic technology to reduce bark beetle

reproduction. Pest Manag Sci 70(1):24–27. https://doi.org/10.1002/ps.3656
Holt MM, Noren DP, Veirs V, Emmons CK, Veirs S (2009) Speaking up: Killer whales (*Orcinus orca*) increase their call amplitude in response to vessel noise. J Acoust Soc Am 125(1):El27–El32. https://doi.org/10.1121/1.3040028
Horn AG, Aikens M, Jamieson E, Kingdon K, Leonard ML (2020) Effect of noise on development of call discrimination by nestling tree swallows, *Tachycineta bicolor*. Anim Behav 164:143–148. https://doi.org/10.1016/j.anbehav.2020.04.008
Houser DS, Martin SW, Finneran JJ (2013a) Behavioural responses of California sea lions to mid-frequency (3250-3450 Hz) sonar signals. Mar Environ Res 92: 268–278. https://doi.org/10.1016/j.marenvres.2013.10.007
Houser DS, Martin SW, Finneran JJ (2013b) Exposure amplitude and repetition affect bottlenose dolphin behavioural responses to simulated mid-frequency sonar signals. J Exp Mar Biol Ecol 443:123–133. https://doi.org/10.1016/j.jembe.2013.02.043
Hu B (2012) Noise-induced structural damage to the cochlea. In: Le Prell CG, Henderson D, Fay RR, Popper AN (eds) Noise-induced hearing loss: scientific advances. Springer, New York, pp 57–86. https://doi.org/10.1007/978-1-4419-9523-0_5
Hu BH, Guo W, Wang PY, Henderson D, Jiang SC (2000) Intense noise-induced apoptosis in hair cells of guinea pig cochleae. Acta Otolaryngol 120(1):19–24. https://doi.org/10.1080/000164800750044443
Hulse SH, MacDougall-Shackleton SA, Wisniewski AB (1997) Auditory scene analysis by songbirds: stream segregation of birdsong by European starlings (*Sturnus vulgaris*). J Comp Psychol 111(1):3–13. https://doi.org/10.1037/0735-7036.111.1.3
International Organization for Standardization (2017) Underwater acoustics—terminology (ISO 18405). Switzerland, Geneva
Jeffrey ML, Brooke MC-E, Douglas HC (2002) Sound detection in situ by the larvae of a coral-reef damselfish (Pomacentridae). Mar Ecol Prog Ser 232:259–268
Jepson PD, Arbelo M, Deaville R, Patterson IAP, Castro P, Baker JR, Degollada E, Ross HM, Herráez P, Pocknell AM, Rodríguez F, Howie FE, Espinosa A, Reid RJ, Jaber JR, Martin V, Cunningham AA, Fernández A (2003) Gas-bubble lesions in stranded cetaceans. Nature 425:575. https://doi.org/10.1038/425575a
Johnston DW (2002) The effect of acoustic harassment devices on harbour porpoises (*Phocoena phocoena*) in the Bay of Fundy, Canada. Biol Conserv 108(1): 113–118. https://doi.org/10.1016/s0006-3207(02)00099-x
Joy R, Tollit D, Wood J, MacGillivray A, Li Z, Trounce K, Robinson O (2019) Potential benefits of vessel slowdowns on endangered southern resident killer whales. Front Mar Sci 6(344). https://doi.org/10.3389/fmars.2019.00344
Juárez R, Araya-Ajoy YG, Barrantes G, Sandoval L (2021) House wrens *Troglodytes aedon* reduce repertoire size and change song element frequencies in response to anthropogenic noise. Ibis 163(1):52–64. https://doi.org/10.1111/ibi.12844
Kaifu K, Akamatsu T, Segawa S (2008) Underwater sound detection by cephalopod statocyst. Fish Sci 74(4): 781–786. https://doi.org/10.1111/j.1444-2906.2008.01589.x
Kastelein RA, Steen N, de Jong C, Wensveen PJ, Verboom WC (2011) Effect of broadband-noise masking on the behavioral response of a harbor porpoise (*Phocoena phocoena*) to 1-s duration 6-7 kHz sonar up-sweeps. J Acoust Soc Am 129(4):2307–2315. https://doi.org/10.1121/1.3559679
Kastelein R, Gransier R, Hoek L, Olthuis J (2012) Temporary threshold shifts and recovery in a harbour porpoise (*Phocoena phocoena*) after octave-band noise at 4 kHz. J Acoust Soc Am 132(5):3525–3537. https://doi.org/10.1121/1.4757641
Kastelein RA, Gransier R, van den Hoogen M, Hoek L (2013) Brief behavioural response threshold levels of a harbour porpoise (*Phocoena phocoena*) to five helicopter dipping sonar signals (1.33 to 1.43 kHz). Aquat Mamm 39(2):162–173
Kerth G, Melber M (2009) Species-specific barrier effects of a motorway on the habitat use of two threatened forest-living bat species. Biol Conserv 142(2): 270–279. https://doi.org/10.1016/j.biocon.2008.10.022
Kight CR, Swaddle JP (2011) How and why environmental noise impacts animals: an integrative, mechanistic review. Ecol Lett 14(10):1052–1061. https://doi.org/10.1111/j.1461-0248.2011.01664.x
Kight CR, Saha MS, Swaddle JP (2012) Anthropogenic noise is associated with reductions in the productivity of breeding Eastern Bluebirds (*Sialia sialis*). Ecol Appl 22(7):1989–1996. https://doi.org/10.1890/12-0133.1
Kobayasi KI, Okanoya K (2003) Context-dependent song amplitude control in Bengalese finches. Neuroreport 14(3):521–524
Krausman PR, Hervert JJ (1983) Mountain sheep responses to aerial surveys. Wildl Soc Bull 11(4): 372–375
Kryter KD (1994) The handbook of hearing and the effects of noise: physiology, psychology, and public health. Academic Press, New York
Kujawa SG, Liberman MC (2009) Adding insult to injury: cochlear nerve degeneration after "temporary" noise-induced hearing loss. J Neurosci 29(45):14077–14085. https://doi.org/10.1523/JNEUROSCI.2845-09.2009
Lampe U, Schmoll T, Franzke A, Reinhold K (2012) Staying tuned: grasshoppers from noisy roadside habitats produce courtship signals with elevated frequency components. Funct Ecol 26(6):1348–1354. https://doi.org/10.1111/1365-2435.12000
Lampe U, Reinhold K, Schmoll T (2014) How grasshoppers respond to road noise: developmental plasticity and population differentiation in acoustic signalling. Funct Ecol 28(3):660–668. https://doi.org/10.1111/1365-2435.12215
Landon DM, Krausman PR, Kiana KGK, Harris LK (2003) Pronghorn use of areas with varying sound pressure levels. Southwest Nat 48(4):725–728

Le Prell CG, Hammill TL, Murphy WJ (2019) Noise-induced hearing loss: translating risk from animal models to real-world environments. J Acoust Soc Am 146(5):3646–3651. https://doi.org/10.1121/1.5133385

Lee N, Mason AC (2017) How spatial release from masking may fail to function in a highly directional auditory system. eLife 6:e20731. https://doi.org/10.7554/eLife.20731

Lee-Dadswell GR (2011) Physics of the interaction between a crab and a seismic test pulse - stage 4: continued development of mathematical model and testing of model via simulation. Cape Breton University, Sydney, NS

Lengagne T (2008) Traffic noise affects communication behaviour in a breeding anuran, *Hyla arborea*. Biol Conserv 141(8):2023–2031. https://doi.org/10.1016/j.biocon.2008.05.017

Leonard ML, Horn AG (2008) Does ambient noise affect growth and begging call structure in nestling birds? Behav Ecol 19(3):502–507. https://doi.org/10.1093/beheco/arm161

Lewis ER, Narins PM (1985) Do frogs communicate with seismic signals? Science 227(4683):187–189. https://doi.org/10.1126/science.227.4683.187

Liberman MC (2016) Noise-induced hearing loss: permanent versus temporary threshold shifts and the effects of hair cell versus neuronal degeneration. In: Popper AN, Hawkins A (eds) The effects of noise on aquatic life II. Springer, New York, pp 1–7. https://doi.org/10.1007/978-1-4939-2981-8_1

Lohr B, Wright TF, Dooling RJ (2003) Detection and discrimination of natural calls in masking noise by birds: estimating the active space signal. Anim Behav 65:763–777

Love EK, Bee MA (2010) An experimental test of noise-dependent voice amplitude regulation in Cope's grey treefrog, *Hyla chrysoscelis*. Anim Behav 80(3):509–515. https://doi.org/10.1016/j.anbehav.2010.05.031

Lovell JM, Findlay MM, Moate RM, Yan HY (2005) The hearing abilities of the prawn *Palaemon serratus*. Comp Biochem Physiol A Mol Integr Physiol 140(1):89–100. https://doi.org/10.1016/j.cbpb.2004.11.003

Lowry H, Lill A, Wong BBM (2012) How noisy does a noisy miner have to be? Amplitude adjustments of alarm calls in an avian urban 'adapter'. PLoS One 7(1):e29960. https://doi.org/10.1371/journal.pone.0029960

Lucke K, Siebert U, Lepper PA, Blanchet M-A (2009) Temporary shift in masked hearing thresholds in a harbor porpoise (*Phocoena phocoena*) after exposure to seismic airgun stimuli. J Acoust Soc Am 125(6):4060–4070. https://doi.org/10.1121/1.3117443

Madsen PT, Wahlberg M, Tougaard J, Lucke K, Tyack P (2006) Wind turbine underwater noise and marine mammals: implications of current knowledge and data needs. Mar Ecol Prog Ser 309:279–295. https://doi.org/10.3354/meps309279

Malme CI, Miles PR, Clark CW, Tyack P, Bird JE (1983) Investigations of the potential effects of underwater noise from petroleum industry activities on migrating gray whale behavior Bolt Beranek and Newman Inc. for U.S. Minerals Management Service, Anchorage, AK, Cambridge, MA

Malme CI, Miles PR, Clark CW, Tyack P, Bird JE (1984) Investigations of the potential effects of underwater noise from petroleum industry activities on migrating gray whale behavior. Phase II: January 1984 migration. (trans: Minerals Management Service USDotI, Washington, DC.). Bolt Beranek and Newman Inc., Anchorage, AK

Manabe K, Sadr EI, Dooling RJ (1998) Control of vocal intensity in budgerigars (*Melopsittacus undulatus*): differential reinforcement of vocal intensity and the Lombard effect. J Acoust Soc Am 103(2):1190–1198. https://doi.org/10.1121/1.421227

Mason JT, McClure CJW, Barber JR (2016) Anthropogenic noise impairs owl hunting behavior. Biol Conserv 199:29–32. https://doi.org/10.1016/j.biocon.2016.04.009

McCarthy E, Moretti D, Thomas L, DiMarzio N, Morrissey R, Jarvis S, Ward J, Izzi A, Dilley A (2011) Changes in spatial and temporal distribution and vocal behavior of Blainville's beaked whales (*Mesoplodon densirostris*) during multiship exercises with mid-frequency sonar. Mar Mamm Sci 27(3):E206–E226. https://doi.org/10.1111/j.1748-7692.2010.00457.x

McCauley RD (2014) Maxima seismic survey noise exposure, Scott Reef 2007. Centre for Marine Science & Technology, Perth, WA

McCauley RD, Fewtrell J, Duncan AJ, Jenner C, Jenner M-N, Penrose JD, Prince RIT, Adhitya A, Murdoch J, McCabe K (2003a) Marine seismic surveys: analysis and propagation of air-gun signals; and effects of exposure on humpback whales, sea turtles, fishes and squid. In: Environmental implications of offshore oil and gas development in Australia: further research. Australian Petroleum Production and Exploration Association, Canberra, ACT

McCauley RD, Fewtrell J, Popper AN (2003b) High intensity anthropogenic sound damages fish ears. J Acoust Soc Am 113:638–642. https://doi.org/10.1121/1.1527962

McCauley RD, Day RD, Swadling KM, Fitzgibbon QP, Watson RA, Semmens JM (2017) Widely used marine seismic survey air gun operations negatively impact zooplankton. Nat Ecol Evol 1:0195. https://doi.org/10.1038/s41559-017-0195

McCue MP, Guinan JJ (1994) Acoustically responsive fibers in the vestibular nerve of the cat. J Neurosci 14(10):6058. https://doi.org/10.1523/JNEUROSCI.14-10-06058.1994

McGregor RL, Bender DJ, Fahrig L (2008) Do small mammals avoid roads because of the traffic? J Appl Ecol 45(1):117–123. https://doi.org/10.1111/j.1365-2664.2007.01403.x

McIntyre E, Leonard ML, Horn AG (2014) Ambient noise and parental communication of predation risk in tree swallows, *Tachycineta bicolor*. Anim Behav 87:85–89. https://doi.org/10.1016/j.anbehav.2013.10.013

McMullen H, Schmidt R, Kunc HP (2014) Anthropogenic noise affects vocal interactions. Behav Process 103:125–128. https://doi.org/10.1016/j.beproc.2013.12.001

Megela-Simmons A, Moss CF, Daniel KM (1985) Behavioral audiograms of the bullfrog (*Rana catesbeiana*) and the green tree frog (*Hyla cinerea*). J Acoust Soc Am 78(4):1236–1244. https://doi.org/10.1121/1.392892

Miller PJO, Biassoni N, Samuels A, Tyack PL (2000) Whale songs lengthen in response to sonar. Nature 405(6789):903. https://doi.org/10.1038/35016148

Miller GW, Moulton VD, Davis RA, Holst M, Millman P, MacGillivray A, Hannay D (2005) Monitoring seismic effects on marine mammals - Southeastern Beaufort Sea, 2001-2002. In: Armsworthy SL, Cranford PJ, Lee K (eds) Offshore oil and gas environmental effects monitoring: approaches and technologies. Battelle Press, Columbus, OH, pp 511–542

Miller PJ, Kvadsheim PH, Lam F-PA, Wensveen PJ, Antunes R, Alves AC, Visser F, Kleivane L, Tyack PL, Sivle LD (2012) The severity of behavioral changes observed during experimental exposures of killer (*Orcinus orca*), long-finned pilot (*Globicephala melas*), and sperm (*Physeter macrocephalus*) whales to naval sonar. Aquat Mamm 38(4):362–401

Miller PJO, Antunes RN, Wensveen PJ, Samarra FIP, Alves AC, Tyack PL, Kvadsheim PH, Kleivane L, Lam F-PA, Ainslie MA, Thomas L (2014) Dose-response relationships for the onset of avoidance of sonar by free-ranging killer whales. J Acoust Soc Am 135(1):975. https://doi.org/10.1121/1.4861346

Miller PJO, Kvadsheim PH, Lam F-PA, Tyack PL, Cure C, Deruiter SL, Kleivane L, Sivle LD, Van Ijsselmuide SP, Visser F, Wensveen PJ, Von Benda-Beckmann AM, Martin Lopez LM, Narazaki T, Hooker SK (2015) First indications that northern bottlenose whales are sensitive to behavioural disturbance from anthropogenic noise. R Soc Open Sci 2:140484. https://doi.org/10.1098/rsos.140484

Mooney TA, Hanlon RT, Christensen-Dalsgaard J, Madsen PT, Ketten DR, Nachtigall PE (2010) Sound detection by the longfin squid (*Loligo pealeii*) studied with auditory evoked potentials: sensitivity to low-frequency particle motion and not pressure. J Exp Biol 213(21):3748–3759. https://doi.org/10.1242/jeb.048348

Mooney TA, Yamato M, Branstetter BK (2012) Hearing in cetaceans: from natural history to experimental biology. Adv Mar Biol 63:197–246

Moore BCJ (2013) An introduction to the psychology of hearing. Brill, Leiden, The Netherlands

Moretti D, Thomas L, Marques T, Harwood J, Dilley A, Neales B, Shaffer J, McCarthy E, New L, Jarvis S, Morrissey R (2014) A risk function for behavioral disruption of Blainville's beaked whales (*Mesoplodon densirostris*) from mid-frequency active sonar. PLoS One 9(1):e85064. https://doi.org/10.1371/journal.pone.0085064

Morley EL, Jones G, Radford AN (2014) The importance of invertebrates when considering the impacts of anthropogenic noise. Proc Biol Sci 281(1776):20132683. https://doi.org/10.1098/rspb.2013.2683

Morris-Drake A, Kern JM, Radford AN (2016) Cross-modal impacts of anthropogenic noise on information use. Curr Biol 26(20):R911–R912. https://doi.org/10.1016/j.cub.2016.08.064

Morton AB, Symonds HK (2002) Displacement of *Orcinus orca* (L.) by high amplitude sound in British Columbia, Canada. ICES J Mar Sci 59(1):71–80

Moseley DL, Derryberry GE, Phillips JN, Danner JE, Danner RM, Luther DA, Derryberry EP (2018) Acoustic adaptation to city noise through vocal learning by a songbird. Proc R Soc B Biol Sci 285(1888):20181356. https://doi.org/10.1098/rspb.2018.1356

Mrosovsky N (1972) Spectrographs of the sounds of leatherback turtles. Herpetologica 28(3):256–258

Narayan R, Best V, Ozmeral E, McClaine E, Dent M, Shinn-Cunningham B, Sen K (2007) Cortical interference effects in the cocktail party problem. Nat Neurosci 10(12):1601–1607. https://doi.org/10.1038/nn2009

Narins PM, Wilson M, Mann DA (2014) Ultrasound detection in fishes and frogs: discovery and mechanisms. In: Köppl C, Manley GA, Popper AN, Fay RR (eds) Insights from comparative hearing research. Springer, New York, pp 133–156. https://doi.org/10.1007/2506_2013_29

National Marine Fisheries Service (2018) 2018 Revisions to: technical guidance for assessing the effects of anthropogenic sound on marine mammal hearing (Version 2.0): underwater thresholds for onset of permanent and temporary threshold shifts. U.-S. Department of Commerce, National Oceanic and Atmospheric Administration, Silver Spring, MD

National Research Council (2005) Marine mammal populations and ocean noise: determining when noise causes biologically significant effects. National Academies Press, Washington. https://doi.org/10.17226/11147

Nemeth E, Pieretti N, Zollinger SA, Geberzahn N, Partecke J, Miranda AC, Brumm H (2013) Bird song and anthropogenic noise: vocal constraints may explain why birds sing higher-frequency songs in cities. Proc Biol Sci 280(1754). https://doi.org/10.1098/rspb.2012.2798

Neo Y, Ufkes E, Kastelein R, Winter H, ten Cate C, Slabbekoorn H (2015) Impulsive sounds change european seabass swimming patterns: influence of pulse repetition interval. Mar Pollut Bull 97(1–2):111–117. https://doi.org/10.1016/j.marpolbul.2015.06.027

New LF, Clark JS, Costa DP, Fleishman E, Hindell MA, Klanjcek T, Lusseau D, Kraus S, McMahon CR,

Robinson PW, Schick RS, Schwarz LK, Simmons SE, Thomas L, Tyack PL, Harwood J (2014) Using short-term measures of behaviour to estimate long-term fitness of southern elephant seals. Mar Ecol Prog Ser 496:99–108. https://doi.org/10.3354/meps10547

Nonaka S, Takahashi R, Enomoto K, Katada A, Unno T (1997) Lombard reflex during PAG-induced vocalization in decerebrate cats. Neurosci Res 29(4):283–289. https://doi.org/10.1016/S0168-0102(97)00097-7

Nordt A, Klenke R (2013) Sleepless in town – drivers of the temporal shift in dawn song in urban European blackbirds. PLoS One 8(8):e71476. https://doi.org/10.1371/journal.pone.0071476

Normandeau Associates (2012) Effects of noise on fish, fisheries, and invertebrates in the U.S. Atlantic and Arctic from energy industry sound-generating activities. U.S. Department of the Interior, Bureau of Ocean Energy Management, Washington, DC

Ocean Studies Board (2016) Approaches to understanding the cumulative effects of stressors on marine mammals. The National Academies Press, Washington, DC. https://doi.org/10.17226/23479

Orci KM, Petróczki K, Barta Z (2016) Instantaneous song modification in response to fluctuating traffic noise in the tree cricket *Oecanthus pellucens*. Anim Behav 112: 187–194. https://doi.org/10.1016/j.anbehav.2015.12.008

Otten W, Kanitz E, Puppe B, Tuchscherer M, Brüssow KP, Nürnberg G, Stabenow B (2004) Acute and long term effects of chronic intermittent noise stress on hypothalamic-pituitary-adrenocortical and sympatho-adrenomedullary axis in pigs. Anim Sci 78(2): 271–283. https://doi.org/10.1017/S1357729800054060

Owen MA, Swaisgood RR, Czekala NM, Steinman K, Lindburg DG (2004) Monitoring stress in captive giant pandas (*Ailuropoda melanoleuca*): behavioral and hormonal responses to ambient noise. Zoo Biol 23(2):147–164. https://doi.org/10.1002/zoo.10124

Parris KM, Velik-Lord M, North JMA (2009) Frogs call at a higher pitch in traffic noise. Ecol Soc 14(1):25

Parry GD, Heislers S, Werner GF, Asplin MD (2002) Assessment of environmental effects of seismic testing on scallop fisheries in Bass Strait. Marine and Freshwater Resources Institute, Queenscliff, VIC

Payne KB, Langbauer WR, Thomas EM (1986) Infrasonic calls of the Asian elephant (*Elephas maximus*). Behav Ecol Sociobiol 18(4):297–301. https://doi.org/10.1007/bf00300007

Payne JF, Andrews CA, Fancey LL, Cook AL, Christian JR (2007) Pilot study on the effects of seismic air gun noise on lobster (*Homarus americanus*). Canadian Technical Report of Fisheries and Aquatic Sciences 2712

Pearson WH, Skalski JR, Malme CI (1992) Effects of Sounds from a Geophysical Survey Device on Behavior of Captive Rockfish (*Sebastes* spp.). Can J Fish Aquat Sci 49(7):1343–1356. https://doi.org/10.1139/f92-150

Penna M, Pottstock H, Velasquez N (2005) Effect of natural and synthetic noise on evoked vocal responses in a frog of the temperate austral forest. Anim Behav 70(3):639–651. https://doi.org/10.1016/j.anbehav.2004.11.022

Pirotta E, Brookes KL, Graham IM, Thompson PM (2014) Variation in harbour porpoise activity in response to seismic survey noise. Biol Lett 10(5):20131090. https://doi.org/10.1098/rsbl.2013.1090

Pohl NU, Slabbekoorn H, Klump GM, Langemann U (2009) Effects of signal features and environmental noise on signal detection in the great tit, *Parus major*. Anim Behav 78(6):1293–1300. https://doi.org/10.1016/j.anbehav.2009.09.005

Pohl NU, Leadbeater E, Slabbekoorn H, Klump GM, Langemann U (2012) Great tits in urban noise benefit from high frequencies in song detection and discrimination. Anim Behav 83(3):711–721. https://doi.org/10.1016/j.anbehav.2011.12.019

Pohl NU, Klump GM, Langemann U (2015) Effects of signal features and background noise on distance cue discrimination by a songbird. J Exp Biol 218(7):1006. https://doi.org/10.1242/jeb.113639

Polajnar J, Čokl A (2008) The effect of vibratory disturbance on sexual behaviour of the southern green stink bug *Nezara viridula* (Heteroptera, Pentatomidae). Cent Eur J Biol 3(2):189–197. https://doi.org/10.2478/s11535-008-0008-7

Polajnar J, Eriksson A, Lucchi A, Anfora G, Virant-Doberlet M, Mazzoni V (2015) Manipulating behaviour with substrate-borne vibrations – potential for insect pest control. Pest Manag Sci 71(1):15–23. https://doi.org/10.1002/ps.3848

Polak M, Wiącek J, Kucharczyk M, Orzechowski R (2013) The effect of road traffic on a breeding community of woodland birds. Eur J For Res 132(5):931–941. https://doi.org/10.1007/s10342-013-0732-z

Popper AN, Fay RR (1993) Sound detection and processing by fish: critical review and major research questions. Brain Behav Evol 41(1):14–25

Popper AN, Fay RR (2011) Rethinking sound detection by fishes. Hear Res 273(1):25–36. https://doi.org/10.1016/j.heares.2009.12.023

Popper AN, Hastings MC (2009) The effects of anthropogenic sources of sound on fishes. J Fish Biol 75:455–489

Popper AN, Hawkins A (eds) (2012) The effects of noise on aquatic life. advances in experimental medicine and biology 730. Springer, New York

Popper AN, Hawkins A (eds) (2016) The effects of noise on aquatic life II. Advances in experimental medicine and biology 730. Springer, New York

Popper AN, Smith ME, Cott PA, Hanna BW, MacGillivray AO, Austin ME, Mann DA (2005) Effects of exposure to seismic airgun use on hearing of three fish species. J Acoust Soc Am 117(6): 3958–3971. https://doi.org/10.1121/1.1904386

Popper AN, Hawkins AD, Fay RR, Mann DA, Bartol S, Carlson TJ, Coombs S, Ellison WT, Gentry RL,

Halvorsen MB, Løkkeborg S, Rogers PH, Southall BL, Zeddies DG, Tavolga WN (2014) Sound exposure guidelines. In: ASA S3/SC1.4 TR-2014 sound exposure guidelines for fishes and sea turtles: a technical report prepared by ANSI-Accredited Standards Committee S3/SC1 and registered with ANSI. Springer, New York, pp 33–51. https://doi.org/10.1007/978-3-319-06659-2_7

Potash LM (1972) Noise-induced changes in calls of the Japanese quail. Psychon Sci 26(5):252–254. https://doi.org/10.3758/bf03328608

Proppe DS, Sturdy CB, St Clair CC (2011) Flexibility in animal signals facilitates adaptation to rapidly changing environments. PLoS One 6(9):e25413. https://doi.org/10.1371/journal.pone.0025413

Przeslawski R, Huang Z, Anderson J, Carroll AG, Edmunds M, Hurt L, Williams S (2018) Multiple field-based methods to assess the potential impacts of seismic surveys on scallops. Mar Pollut Bull 129(2):750–761. https://doi.org/10.1016/j.marpolbul.2017.10.066

Quinn JL, Whittingham MJ, Butler SJ, Cresswell W (2006) Noise, predation risk compensation and vigilance in the chaffinch *Fringilla coelebs*. J Avian Biol 37(6):601–608. https://doi.org/10.1111/j.2006.0908-8857.03781.x

Rabin LA, Coss RG, Owings DH (2006) The effects of wind turbines on antipredator behavior in California ground squirrels (*Spermophilus beecheyi*). Biol Conserv 131(3):410–420. https://doi.org/10.1016/j.biocon.2006.02.016

Raboin M, Elias DO (2019) Anthropogenic noise and the bioacoustics of terrestrial invertebrates. J Exp Biol 222(12):jeb178749. https://doi.org/10.1242/jeb.178749

Ratnam R, Feng AS (1998) Detection of auditory signals by frog inferior collicular neurons in the presence of spatially separated noise. J Neurophysiol 80(6):2848

Reichmuth C, Holt MM, Mulsow J, Sills JM, Southall BL (2013) Comparative assessment of amphibious hearing in pinnipeds. J Comp Physiol A 199(6):491–507

Reijnen R, Foppen R (2006) Impact of road traffic on breeding bird populations. In: Davenport J, Davenport JL (eds) The ecology of transportation: managing mobility for the environment. Springer, Netherlands, Dordrecht, pp 255–274. https://doi.org/10.1007/1-4020-4504-2_12

Reijnen R, Foppen R, Braak CT, Thissen J (1995) The effects of car traffic on breeding bird populations in woodland. III. Reduction of density in relation to the proximity of main roads. J Appl Ecol 32(1):187–202. https://doi.org/10.2307/2404428

Rheindt FE (2003) The impact of roads on birds: does song frequency play a role in determining susceptibility to noise pollution? J Ornithol 144(3):295–306. https://doi.org/10.1007/bf02465629

Richardson WJ, Greene CR, Malme CI, Thomson DH (1995) Marine mammals and noise. Academic Press, San Diego

Richardson AJ, Matear RJ, Lenton A (2017) Potential impacts on zooplankton of seismic surveys. CSIRO, Canberra, Australia

Ridgway SH, Wever EG, McCormick JG, Palin J, Anderson JH (1969) Hearing in the giant sea turtle, *Chelonia mydas*. Proc Natl Acad Sci USA 64:884–890

Roian Egnor SE, Hauser MD (2006) Noise-induced vocal modulation in cotton-top tamarins (*Saguinus oedipus*). Am J Primatol 68(12):1183–1190. https://doi.org/10.1002/ajp.20317

Rolland RM, Parks SE, Hunt KE, Castellote M, Corkeron PJ, Nowacek DP, Wasser SK, Kraus SD (2012) Evidence that ship noise increases stress in right whales. Proc R Soc Lond Ser B Biol Sci 279(1737):2363–2368. https://doi.org/10.1098/rspb.2011.2429

Romano TA, Keogh MJ, Kelly C, Feng P, Berk L, Schlundt CE, Carder DA, Finneran JJ (2004) Anthropogenic sound and marine mammal health: measures of the nervous and immune systems before and after intense sound exposure. Can J Fish Aquat Sci 61(7):1124–1134. https://doi.org/10.1139/f04-055

Romero ML, Butler LK (2007) Endocrinology of stress. Int J Comp Psychol 20(2)

Ruffoli R, Carpi A, Giambelluca MA, Grasso L, Scavuzzo MC, Giannessi FF (2006) Diazepam administration prevents testosterone decrease and lipofuscin accumulation in testis of mouse exposed to chronic noise stress. Andrologia 38(5):159–165. https://doi.org/10.1111/j.1439-0272.2006.00732.x

Ryals BM, Rubel EW (1988) Hair cell regeneration after acoustic trauma in adult *Coturnix* quail. Science 240(4860):1774. https://doi.org/10.1126/science.3381101

Samuel Y, Morreale SJ, Clark CW, Greene CH, Richmond ME (2005) Underwater, low-frequency noise in a coastal sea turtle habitat. J Acoust Soc Am 117(3):1465–1472. https://doi.org/10.1121/1.1847993

Saunders JC, Dooling RJ (2018) Characteristics of temporary and permanent threshold shifts in vertebrates. In: Slabbekoorn H, Dooling RJ, Popper AN, Fay RR (eds) Effects of anthropogenic noise on animals. Springer, New York, pp 83–107. https://doi.org/10.1007/978-1-4939-8574-6_4

Schaub A, Ostwald J, Siemers BM (2008) Foraging bats avoid noise. J Exp Biol 211(19):3174. https://doi.org/10.1242/jeb.022863

Scheifele PM, Andrew S, Cooper RA, Darre M, Musiek FE, Max L (2005) Indication of a Lombard vocal response in the St. Lawrence River beluga. J Acoust Soc Am 117(3):1486–1492. https://doi.org/10.1121/1.1835508

Schmidt AKD, Riede K, Römer H (2011) High background noise shapes selective auditory filters in a tropical cricket. J Exp Biol 214(10):1754. https://doi.org/10.1242/jeb.053819

Schmidt R, Morrison A, Kunc HP (2014) Sexy voices – no choices: male song in noise fails to attract females. Anim Behav 94:55–59. https://doi.org/10.1016/j.anbehav.2014.05.018

Serrano A, Terhune JM (2001) Within-call repetition may be an anti-masking strategy in underwater calls of harp seals. Can J Zool 79:1410–1413

Shannon G, Angeloni LM, Wittemyer G, Fristrup KM, Crooks KR (2014) Road traffic noise modifies behaviour of a keystone species. Anim Behav 94: 135–141. https://doi.org/10.1016/j.anbehav.2014.06.004

Shannon G, Crooks KR, Wittemyer G, Fristrup KM, Angeloni LM (2016) Road noise causes earlier predator detection and flight response in a free-ranging mammal. Behav Ecol 27(5):1370–1375. https://doi.org/10.1093/beheco/arw058

Shen J-X, Feng AS, Xu Z-M, Yu Z-L, Arch VS, Yu X-J, Narins PM (2008) Ultrasonic frogs show hyperacute phonotaxis to female courtship calls. Nature 453(7197):914–916. https://doi.org/10.1038/nature06719

Siemers BM, Schaub A (2011) Hunting at the highway: traffic noise reduces foraging efficiency in acoustic predators. Proc R Soc B Biol Sci 278(1712): 1646–1652. https://doi.org/10.1098/rspb.2010.2262

Simmons AM, Narins PM (2018) Effects of anthropogenic noise on amphibians and reptiles. In: Slabbekoorn H, Dooling RJ, Popper AN, Fay RR (eds) Effects of anthropogenic noise on animals. Springer, New York, pp 179–208. https://doi.org/10.1007/978-1-4939-8574-6_7

Simpson SD, Meekan M, Montgomery J, McCauley R, Jeffs A (2005) Homeward sound. Science 308:221. https://doi.org/10.1126/science.1107406

Simpson SD, Jeffs A, Montgomery JC, McCauley RD, Meekan MG (2007) Nocturnal relocation of adult and juvenile coral reef fishes in response to reef noise. Coral Reefs 27:97–104. https://doi.org/10.1007/s00338-007-0294-y

Simpson SD, Radford AN, Nedelec SL, Ferrari MCO, Chivers DP, McCormick MI, Meekan MG (2016) Anthropogenic noise increases fish mortality by predation. Nat Commun 7:10544. https://doi.org/10.1038/ncomms10544

Slabbekoorn H, Peet M (2003) Ecology: birds sing at a higher pitch in urban noise. Nature 424(6946): 267–267. https://doi.org/10.1038/424267a

Slabbekoorn H, Ripmeester EAP (2008) Birdsong and anthropogenic noise: implications and applications for conservation. Mol Ecol 17(1):72–83. https://doi.org/10.1111/j.1365-294X.2007.03487.x

Slabbekoorn H, Bouton N, van Opzeeland I, Coers A, ten Cate C, Popper AN (2010) A noisy spring: the impact of globally rising underwater sound levels on fish. Trends Ecol Evol 25(7):419–427. https://doi.org/10.1016/j.tree.2010.04.005

Slabbekoorn H, Dooling RJ, Popper AN, Fay RR (2018a) Effects of anthropogenic noise on animals. In: Springer handbook of auditory research, vol 66. Springer, New York

Slabbekoorn H, McGee J, Walsh EJ (2018b) Effects of man-made sound on terrestrial mammals. In: Slabbekoorn H, Dooling RJ, Popper AN, Fay RR (eds) Effects of anthropogenic noise on animals. Springer, New York, pp 243–276. https://doi.org/10.1007/978-1-4939-8574-6_9

Slotte A, Hansen K, Dalen J, Ona E (2004) Acoustic mapping of pelagic fish distribution and abundance in relation to a seismic shooting area off the Norwegian west coast. Fish Res 67(2):143–150. https://doi.org/10.1016/j.fishres.2003.09.046

Sobrian SK, Vaughn VT, Ashe WK, Markovic B, Djuric V, Jankovic BD (1997) Gestational exposure to loud noise alters the development and postnatal responsiveness of humoral and cellular components of the immune system in offspring. Environ Res 73(1):227–241. https://doi.org/10.1006/enrs.1997.3734

Söffker M, Trathan P, Clark J, Collins MA, Belchier M, Scott R (2015) The impact of predation by marine mammals on patagonian toothfish longline fisheries. PLoS One 10(3):e0118113. https://doi.org/10.1371/journal.pone.0118113

Solé M, Lenoir M, Durfort M, López-Bejar M, Lombarte A, André M (2013) Ultrastructural damage of *Loligo vulgaris* and *Illex coindetii* statocysts after low frequency sound exposure. PLoS One 8(10): e78825. https://doi.org/10.1371/journal.pone.0078825

Song J, Mann DA, Cott PA, Hanna BW, Popper AN (2008) The inner ears of Northern Canadian freshwater fishes following exposure to seismic air gun sounds. J Acoust Soc Am 124(2):1360–1366. https://doi.org/10.1121/1.2946702

Southall BL (2018) Noise. In: Würsig B, Thewissen JGM, Kovacs KM (eds) Encyclopedia of marine mammals, 3rd edn. Academic Press, New York, pp 637–645. https://doi.org/10.1016/B978-0-12-804327-1.00183-7

Southall BL, Bowles AE, Ellison WT, Finneran JJ, Gentry RL, Greene CRJ, Kastak D, Ketten DR, Miller JH, Nachtigall PE, Richardson WJ, Thomas JA, Tyack PL (2007) Marine mammal noise exposure criteria: Initial scientific recommendations. Aquat Mamm 33(4):411–521. https://doi.org/10.1080/09524622.2008.9753846

Southall BL, Nowacek DP, Miller PJO, Tyack PL (2016) Experimental field studies to measure behavioral responses of cetaceans to sonar. Endanger Species Res 31:293–315. https://doi.org/10.3354/esr00764

Southall BL, DeRuiter SL, Friedlaender A, Stimpert AK, Goldbogen JA, Hazen E, Casey C, Fregosi S, Cade DE, Allen AN, Harris CM, Schorr G, Moretti D, Guan S, Calambokidis J (2019a) Behavioral responses of individual blue whales (*Balaenoptera musculus*) to mid-frequency military sonar. J Exp Biol 222(5): jeb190637. https://doi.org/10.1242/jeb.190637

Southall BL, Finneran JJ, Reichmuth C, Nachtigall PE, Ketten DR, Bowles AE, Ellison WT, Nowacek DP, Tyack PL (2019b) Marine mammal noise exposure criteria: updated scientific recommendations for residual hearing effects. Aquat Mamm 45(2):125–232. https://doi.org/10.1578/AM.45.2.2019.125

Stanley JA, Radford CA, Jeffs AG (2009) Induction of settlement in crab megalopae by ambient underwater reef sound. Behav Ecol 21(1):113–120. https://doi.org/10.1093/beheco/arp159

Stanley JA, Wilkens SL, Jeffs AG (2014) Fouling in your own nest: vessel noise increases biofouling. Biofouling 30(7):837–844. https://doi.org/10.1080/08927014.2014.938062

Stimpert AK, Deruiter SL, Southall BL, Moretti DJ, Falcone EA, Goldbogen JA, Friedlaender A, Schorr GS, Calambokidis J (2014) Acoustic and foraging behavior of a Baird's beaked whale, *Berardius bairdii*, exposed to simulated sonar. Sci Rep 4:7031. https://doi.org/10.1038/srep07031

Strasser EH, Heath JA (2013) Reproductive failure of a human-tolerant species, the American kestrel, is associated with stress and human disturbance. J Appl Ecol 50(4):912–919. https://doi.org/10.1111/1365-2664.12103

Sun JWC, Narins PM (2005) Anthropogenic sounds differentially affect amphibian call rate. Biol Conserv 121(3):419–427. https://doi.org/10.1016/j.biocon.2004.05.017

Swaddle JP, Page LC (2007) High levels of environmental noise erode pair preferences in zebra finches: implications for noise pollution. Anim Behav 74(3):363–368. https://doi.org/10.1016/j.anbehav.2007.01.004

Tarlow EM, Blumstein DT (2007) Evaluating methods to quantify anthropogenic stressors on wild animals. Appl Anim Behav Sci 102(3):429–451. https://doi.org/10.1016/j.applanim.2006.05.040

Tavolga WN (ed) (1976) Sound reception in fishes. Dowden, Hutchinson and Ross, Stroudsburg, PA

Tavolga WN, Popper AN, Fay RR (eds) (2012) Hearing and sound communication in fishes. Springer, New York

Tennessen JB, Parks SE, Langkilde T (2014) Traffic noise causes physiological stress and impairs breeding migration behaviour in frogs. Conserv Physiol 2(1). https://doi.org/10.1093/conphys/cou032

Tennessen JB, Parks SE, Swierk L, Reinert LK, Holden WM, Rollins-Smith LA, Walsh KA, Langkilde T (2018) Frogs adapt to physiologically costly anthropogenic noise. Proc R Soc B Biol Sci 285(1891):20182194. https://doi.org/10.1098/rspb.2018.2194

Thode AM, Blackwell SB, Conrad AS, Kim KH, Marques T, Thomas L, Oedekoven CS, Harris D, Bröker K (2020) Roaring and repetition: how bowhead whales adjust their call density and source level (Lombard effect) in the presence of natural and seismic airgun survey noise. J Acoust Soc Am 147(3):2061–2080. https://doi.org/10.1121/10.0000935

Thomas JA, Friel B, Yegge S (2016) Restoring dueting behavior in a mated pair of buffy cheeked gibbons after exposure to construction noise at a zoo through playbacks of their own sounds. J Acoust Soc Am 140(4):3415–3415. https://doi.org/10.1121/1.4970975

Thompson PM, Lusseau D, Barton T, Simmons D, Rusin J, Bailey H (2010) Assessing the responses of coastal cetaceans to the construction of offshore wind turbines. Mar Pollut Bull 60(8):1200–1208. https://doi.org/10.1016/j.marpolbul.2010.03.030

Thompson PM, Brookes KL, Graham IM, Barton TR, Needham K, Bradbury G, Merchant ND (2013) Short-term disturbance by a commercial two-dimensional seismic survey does not lead to long-term displacement of harbour porpoises. Proc R Soc Lond Ser B Biol Sci 280(1771). https://doi.org/10.1098/rspb.2013.2001

Thomsen F, Erbe C, Hawkins A, Lepper P, Popper AN, Scholik-Schlomer A, Sisneros J (2020) Introduction to the special issue on the effects of sound on aquatic life. J Acoust Soc Am 148(2):934–938. https://doi.org/10.1121/10.0001725

Todd VLG, Todd IB, Gardiner JC, Morrin ECN, Macpherson NA, Dimarzio NA, Thomsen F (2015) A review of impacts of marine dredging activities on marine mammals. ICES (International Council for the Exploration of the Seas). J Mar Sci 77(2):328–340. https://doi.org/10.1093/icesjms/fsu187

Tougaard J, Carstensen J, Teilmann J (2009) Pile driving zone of responsiveness extends beyond 20 km for harbor porpoises (*Phocoena phocoena* (L.)). J Acoust Soc Am 126(1):11–14. https://doi.org/10.1121/1.3132523

Turnbull SD (1994) Changes in masked thresholds of a harbor seal *Phoca vitulina* associated with angular separation of signal and noise sources. Can J Zool 72:1863–1866. https://doi.org/10.1139/z94-253

Tyack PL, Zimmer WMX, Moretti D, Southall BL, Claridge DE, Durban JW, Clark CW, D'Amico A, DiMarzio N, Jarvis S, McCarthy E, Morrissey R, Ward J, Boyd IL (2011) Beaked whales respond to simulated and actual navy sonar. PLoS One 6(3):e17009. https://doi.org/10.1371/journal.pone.0017009

Valero MD, Hancock KE, Maison SF, Liberman MC (2018) Effects of cochlear synaptopathy on middle-ear muscle reflexes in unanesthetized mice. Hear Res 363:109–118. https://doi.org/10.1016/j.heares.2018.03.012

Vermeij MJA, Marhaver KL, Huijubers C, Nagelkerken I, Simpson SD (2010) Coral larvae move toward reef sounds. PLoS One 5(5):e10660. https://doi.org/10.1371/journal.pone.0010660

Verzijden MN, Ripmeester EAP, Ohms VR, Snelderwaard P, Slabbekoorn H (2010) Immediate spectral flexibility in singing chiffchaffs during experimental exposure to highway noise. J Exp Biol 213(15):2575. https://doi.org/10.1242/jeb.038299

Visser F, Curé C, Kvadsheim PH, Lam F-PA, Tyack PL, Miller PJO (2016) Disturbance-specific social responses in long-finned pilot whales, *Globicephala melas*. Sci Rep. https://doi.org/10.1038/srep28641

Wale MA, Simpson SD, Radford AN (2013a) Noise negatively affects foraging and antipredator behaviour in

shore crabs. Anim Behav 86(1):111–118. https://doi.org/10.1016/j.anbehav.2013.05.001

Wale MA, Simpson SD, Radford AN (2013b) Size-dependent physiological responses of shore crabs to single and repeated playback of ship noise. Biol Lett 9(2):20121194. https://doi.org/10.1098/rsbl.2012.1194

Ward AI, Pietravalle S, Cowan DP, Delahay RJ (2008) Deterrent or dinner bell? Alteration of badger activity and feeding at baited plots using ultrasonic and water jet devices. Appl Anim Behav Sci 115(3):221–232. https://doi.org/10.1016/j.applanim.2008.06.004

Warnecke M, Chiu C, Engelberg J, Moss CF (2015) Active listening in a bat cocktail party: adaptive echolocation and flight behaviors of big brown bats, *Eptesicus fuscus*, foraging in a cluttered acoustic environment. Brain Behav Evol 86(1):6–16. https://doi.org/10.1159/000437346

Warren B, Fenton GE, Klenschi E, Windmill JFC, French AS (2020) Physiological basis of noise-induced hearing loss in a tympanal ear. J Neurosci 40(15):3130. https://doi.org/10.1523/JNEUROSCI.2279-19.2019

Weir CR, Dolman SJ (2007) Comparative review of the regional marine mammal mitigation guidelines implemented during industrial seismic surveys, and guidance towards a worldwide standard. J Int Wildl Law Policy 10:1–27. https://doi.org/10.1080/13880290701229838

Weisenberger ME, Krausman PR, Wallace MC, De Young DW, Maughan OE (1996) Effects of simulated jet aircraft noise on heart rate and behavior of desert ungulates. J Wildl Manag 60(1):52–61. https://doi.org/10.2307/3802039

Wensveen PJ, Kvadsheim PH, Lam F-PA, von Benda-Beckmann AM, Sivle LD, Visser F, Curé C, Tyack PL, Miller PJO (2017) Lack of behavioural responses of humpback whales (*Megaptera novaeangliae*) indicate limited effectiveness of sonar mitigation. J Exp Biol 220(22):4150–4161. https://doi.org/10.1242/jeb.161232

Wever EG (1978) The reptile ear. Princeton University Press, Princeton. https://doi.org/10.2307/j.ctvbcd2f0

Wever EG (1985) The amphibian ear. Princeton University Press. https://doi.org/10.2307/j.ctt7zth8g

Williams R, Erbe C, Ashe E, Beerman A, Smith J (2014) Severity of killer whale behavioural responses to ship noise: a dose-response study. Mar Pollut Bull 79:254–260. https://doi.org/10.1016/j.marpolbul.2013.12.004

Williams R, Wright AJ, Ashe E, Blight LK, Bruintjes R, Canessa R, Clark CW, Cullis-Suzuki S, Dakin DT, Erbe C, Hammond PS, Merchant ND, O'Hara PD, Purser J, Radford AN, Simpson SD, Thomas L, Wale MA (2015) Impacts of anthropogenic noise on marine life: publication patterns, new discoveries, and future directions in research and management. Ocean Coast Manag 115:17–24. https://doi.org/10.1016/j.ocecoaman.2015.05.021

Wisniewska DM, Johnson M, Teilmann J, Rojano-Doñate L, Shearer J, Sveegaard S, Miller Lee A, Siebert U, Madsen Peter T (2016) Ultra-high foraging rates of harbor porpoises make them vulnerable to anthropogenic disturbance. Curr Biol 26(11):1441–1446. https://doi.org/10.1016/j.cub.2016.03.069

Wollerman L, Wiley RH (2002) Background noise from a natural chorus alters female discrimination of male calls in a neotropical frog. Anim Behav 63(1):15–22. https://doi.org/10.1006/anbe.2001.1885

World Health Organization (2011) Burden of disease from environmental noise: quantification of healthy life years lost in Europe. World Health Organization, Copenhagen

Wrege PH, Rowland ED, Thompson BG, Batruch N (2010) Use of acoustic tools to reveal otherwise cryptic responses of forest elephants to oil exploration. Conserv Biol 24(6):1578–1585. https://doi.org/10.1111/j.1523-1739.2010.01559.x

Yelverton JT, Richmond DR, Hicks W, Saunders K, Fletcher ER (1975) The relationship between fish size and their response to underwater blast. Lovelace Foundation for Medical Education and Research, Albuquerque, NM

Yi YZ, Sheridan JA (2019) Effects of traffic noise on vocalisations of the rhacophorid tree frog *Kurixalus chaseni* (Anura: Rhacophoridae) in Borneo. RAFFLES Bull Zool 67:77–82. https://doi.org/10.26107/RBZ-2019-0007

Young BA (1997) A review of sound production and hearing in snakes, with a discussion of intraspecific acoustic communication in snakes. J Pa Acad Sci 71(1):39–46

Zhao L, Zhu B, Wang J, Brauth SE, Tang Y, Cui J (2017) Sometimes noise is beneficial: stream noise informs vocal communication in the little torrent frog *Amolops torrentis*. J Ethol 35(3):259–267. https://doi.org/10.1007/s10164-017-0515-y

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



# Index


**A**
Abiotic noise, 175, 459
Absolute abundance, 348
Absorption, 160, 161, 187, 194, 195
Accelerometer, 18
Acoustic adaptation hypothesis, 218, 236
Acoustic alarms, 234
Acoustically hard, 165
Acoustically soft, 165
Acoustic camera, 17, 43
Acoustic communication, 187
Acoustic complexity index (ACI), 248
Acoustic diversity index (ADI), 247
Acoustic energy, 119
Acoustic environment, 217
Acoustic evenness index (AEI), 247
Acoustic habitat generalists, 218, 219
Acoustic habitat hypothesis, 218
Acoustic habitat specialists, 218, 219
Acoustic harassment devices, 465
Acoustic impedance, 195, 199, 422
Acoustic indices, 246
Acoustic intensity, 120
Acoustic masking, 486
Acoustic mirage, 166, 169
Acoustic niche hypothesis, 218, 219
Acoustic niches, 242
Acoustic power, 120
Acoustic scene, 219, 427
Acoustic startle response, 360
Acoustic tag, 26
Acoustic trauma, 462
Acoustic wavefront, 193
Active sonar, 21
Active sonar systems, 199
Active space, 177, 475, 482, 491
AD-converters, 43
Adiabatic mode method, 205
Advertisement calls, 363
Affected source level, 126
Affiliative calls, 404
African elephant, 221, 296
Aggressive calls, 404
Air absorption, 167
Aircrafts, 225
Airgun, 210, 238, 249, 468–470
Airplane, 226
Akaike's information criterion (AIC), 346
Alarm calls, 98, 406
Alder flycatchers, 173
Aliasing, 137
Alligator, 222
Alpacas, 366
Alternative hypothesis, 337
Altruism, 406
Ambient noise, 171, 187, 357
Ambient sound, 228
American bullfrog, 362
Amplitude, 116
Amplitude-modulation, 116
Amplitude sensitivity, 39
Analog, 2, 137
Analog-to-digital converter, 3
ANCOVA, 341
Anechoic chamber, 360
Angle of incidence, 195
Angular frequency, 200
Animal choruses, 278
Animal ethics permits, 88
ANOVA, 341
Antagonistic, 463
Antarctic blue whales, 236, 304
Antarctic minke whale, 273
Antennae, 471
Anthropogenic noise, 175, 460
Anthropophony, 112, 218
Anti-aliasing filter, 138
Anti-masking strategies, 172, 187
Antinode, 203
Antiphonal, 392
Ants, 220
Anuran, 220, 295, 296
Appraiser, 393
Approach calls, 423
Aquatic biophony, 228
Aquatic geophony, 232
Aquatic soundscape, 227
Aquatic technophony, 233


© The Author(s) 2022
C. Erbe, J. A. Thomas (eds.), *Exploring Animal Behavior Through Sound: Volume 1*,
https://doi.org/10.1007/978-3-030-97540-1

Arctic fox, 371
Artificial neural networks, 291
Atmospheric pressure, 116
Atui swiftlet, 442
Audibility, 460
Audience effect, 400
Audio coding, 139
Audiogram, 187, 356
Auditory brainstem response, 357, 375
Auditory evoked potentials, 358
Auditory nerve, 361
Auditory scene analysis, 219
Auditory-vestibular nerve, 361
Australian grey swiftlet, 441
Australian sea lion, 223
Autocorrelated, 331
Autonomous recorder, 13
Autonomous underwater vehicles (AUVs), 25
A-weighting, 129

**B**
Backpropagation, 293
Badgers, 486
Band-pass filter, 131
Bandwidth, 38, 40, 131
Barbastelle bats, 92
Barnacles, 460
Barred owl, 271
Bat, 160, 172, 222, 238, 271, 294, 300, 485
Bat detectors, 19, 43
Bayesian inference, 335
Bayesian information criterion (BIC), 346
Beaked whales, 490
Beamforming, 22, 55, 147
Bearded seals, 98
Behavioral hearing threshold, 357
Behavioral plasticity, 472
Behavioral response, 465
Beluga whale, 187, 188, 239, 359, 367, 371, 372, 466
Best hearing sensitivity, 358
Beta distribution, 327
Bias, 323
Bidirectional, 44
Big brown bat, 430, 460
Bimodal acousto-vibrational communication, 398
Binary coding, 3
Binaural, 46
Binomial distribution, 327
Bioacoustic index, 246
Biofouling, 470
Biophony, 112, 218
Biosonar, 193, 222
Biotic noise, 175, 459
Biotremology, 18, 390
Birds, 168, 219, 295
Bit depth, 3, 138
Bivariate, 332
Blackbird, 163
Black-capped chickadee, 292, 430, 483, 484
Black lemur, 286
Black-tufted marmosets, 486
Blainville's beaked whales, 290, 295, 297
Blue whales, 278, 489
Blur-ratio, 162
Boat, 226
Bootstrap, 329
Bornean rock frog, 476
Bottlenose dolphin, 102, 284, 297, 300, 372, 378, 379, 432, 439, 490
Bowhead whale, 462
Brazilian free-tailed bat, 486
Broadband, 131
Brown long-eared bat, 429
Bryozoans, 460
Bubbles, 199
Budgerigar, 178, 360, 476, 482, 484
Bullfrogs, 176
Bumble bees, 220
Buntings, 297
Burst-pulse sounds, 231
Butterfly, 473

**C**
Calibration signal, 90
California sea lion, 372, 490
Campbell's monkeys, 271
Canaries, 482
Car, 225, 226
Cardioid, 44
Catch trials, 366
Categorical variables, 326
Cats, 486, 487
Caustics, 201
Censored data, 343
Center frequency, 131
Central tendency, 326
Cephalopods, 469
Cepstral coefficients, 283
Cepstrogram, 284
Chaffinches, 466, 483
Characteristic impedance, 121
Chickens, 484
Chiffchaff, 220
Chimpanzees, 173
Chinchilla, 356, 487
Chipmunks, 271, 288, 465
Chorus, 482
Cicadas, 220
Cichlids, 96
Classical conditioning, 365
Classification tree, 284, 288
Clicks, 231
Closest point of approach, 142, 197
Cochlea, 485
Cochlear microphonic potential, 375
Cockatoos, 302
Cocktail party effect, 401
Coefficient of variation, 329

Coherent, 124, 198
Comb-filter effects, 165
Common chiffchaffs, 225
Common dolphin, 231, 284
Common eastern froglet, 475
Common marmosets, 486
Common octopus, 377
Common yellowthroat, 88
Communication, 390, 459
Communication space, 224, 491
Comodulation masking release, 172, 178, 491
Compact cassette, 8
Compound action potential, 375
Compressional waves, 199
Compressions, 113
Condenser microphone, 15, 42
Conditioned response, 102, 358, 359, 365
Conditioned stimulus, 365
Conditioned suppression, 366
Confidence intervals, 337
Confounding effects, 326
Confounding factors, 326
Confusion matrix, 280
Conservative response bias, 371
Construction equipment, 226
Constructive interference, 164
Context-dependent, 403
Context-independent, 403, 410
Continuous, 326
Continuous variables, 326
Continuous wave, 115
Contour, 115
Convective air currents, 171
Convergence zone, 202, 235
Convergent evolution, 423
Convolutional neural networks, 293
Copepod, 468, 471
Coqui frogs, 239
Corals, 468, 470
Correct detection, 370
Correct rejection, 370
Cortical evoked responses, 357
Coterie, 406
Cotton-top tamarins, 486
Coupled mode method, 205
Covariates, 326
Crab, 229, 465, 468, 470
Crabeater seals, 93
Credible intervals, 337
Crepuscular, 401
Crickets, 178, 220, 466, 472
Critical bandwidth, 382
Critical ratio, 172, 177, 187, 381, 461, 477
Crocodiles, 220
Cross-correlation, 144
Cross-modal impacts, 467
Crustaceans, 467, 468
CTD measurements, 101
Cultural trait, 402
Curvature, 115
Cut-off frequency, 131
Cuttlefish, 468, 469
C-weighting, 129
Cylindrical spreading, 159, 193, 195

**D**
Damselfish, 96
Datasheet, 94
Daubenton's bat, 421
3D audio, 17
3-dB bandwidth, 131
10-dB bandwidth, 131
3-dB duration, 132
10-dB duration, 132
Dear enemy effect, 173
Decibels, 123
Decidecades, 132
Deep isothermal layer, 190
Deep neural networks, 293
Deep sound channel (DSC), 202
Dendrogram, 284
Dependent variables, 326
Depth sensor, 55
Descriptive studies, 321
Desensitization, 473
Desert mule deer, 486
Desert tortoises, 474
Designators, 393
Destructive interference, 164
Detection threshold, 187, 276
Developmental plasticity, 472
Dialects, 402
Difference limen, 379
Diffraction, 162, 171, 201
Diffractive scattering, 163
Diffuse reflection, 161
Digital, 3
Digital audio tape, 10
Digital compact cassette, 10
Digital recorders, 10, 43
Digital-to-analog converter, 3
Dinner bell effect, 460, 486
Dipole sound field, 478
Directionality, 42
Directional microphone, 17
Directivity index, 187
Discrete variables, 326
Discriminant function analysis, 287
Displacement from habitat, 466
Distant shipping, 228
Distortion product otoacoustic emissions, 373
Distracted prey hypothesis, 467
Distress calls, 406
Dogs, 224
Domain, 327
Domestic kittens, 365
Doppler shift, 142, 226, 428
Doppler-shift compensators, 428

Dose-response curve, 462, 489, 490
Downsweep, 115
Downward refraction, 158, 166, 168, 169
Downwind, 169
Drumming, 396
Duck, 220
Duct, 201
Duetting, 98, 408
Dusk chorus, 409
Duty cycle, 427
Dwarf minke whale, 273
Dynamic information, 403
Dynamic microphones, 15, 42
Dynamic range, 39, 40
Dynamic range of hearing, 380
Dynamic time-warping, 297

**E**
Earthquakes, 232, 233
Eastern bluebirds, 480
Eavesdropping, 400, 446
Echolocation, 19, 222
Ecoacoustics, 218
Ecuadorian hillstar hummingbird, 220
Eddies, 171
Effect size, 333
Egrets, 466
Egyptian fruit bat, 98, 430
Elastic seabeds, 199
Electret condenser microphones, 16
Electret microphones, 42
Electrophysiological hearing threshold, 357
Electrophysiological noise, 357
Electrostatic microphones, 42
Elephant, 168, 174, 239, 250, 485
Elephant seal, 463
Elk, 486
Empirical model, 329
Emulation, 392
End frequency, 115
Energetic masking, 461
90% energy signal duration, 119, 211
Energy threshold detector, 278
Entropy index, 247
Environmental conditions, 91
Equal loudness contours, 129
Equal-power assumption, 188
Error of anticipation, 369
Error of habituation, 369
Estimators, 329
European blackbirds, 483, 484
European robins, 483, 484
Evanescent wave, 143
Evoked calling, 365
Ewes, 224
Excess attenuation, 160
Explanatory variables, 326
Exploratory studies, 321
Explosions, 234
Extended source, 127
External auditory meatus, 358
Extrapolation, 325

**F**
Fallow deer, 173
False alarm, 280, 366, 370
False killer whale, 101, 304, 432
False negative, 339
False positive, 280, 339
Far-field, 126, 127
Fast-field programs, 205
Fear-potentiated startle, 362
Feature vectors, 282
Field crickets, 472
Field quantity, 123
File formats, 104
Filter, 131
Fin whale, 208, 236, 278
Fish, 229, 245, 284, 466, 477
Fish chorus, 230, 243
Fixed effects, 342
Fletcher critical band, 382
Flies, 471
Foliage attenuation, 165
Foot-flagging frogs, 466
Forest canopy, 168
Forest elephants, 486
Formant dispersion, 404
Formants, 174
Forward masking, 383
Fourier transform, 136
Free-field, 126, 127
Frequency, 113, 356
Frequency band, 131
Frequency dispersion, 300
Frequency-division, 60
Frequency domain, 134, 136
Frequency domain models, 200
Frequency-modulation, 115, 420
Frequency region of best sensitivity, 358
Frequency resolution, 141
Frequency response, 38, 40
Frequency selectivity, 380
Frequency spacing, 139, 141
Frequency weightings, 128
Frequentist inference, 335
Frog, 220, 474
Fruit fly, 176
Functionally referential signals, 406
Fundamental frequency, 114, 220

**G**
Gabor signals, 116
Gamma distribution, 327
Gated recurrent unit, 293
Gaussian beam tracing, 201, 207
Gaussian distribution, 327
Gaussian linear regression model, 330

Gaussian mixture models (GMMs), 295
Generalized additive mixed models (GAMM), 342
Generalized additive model (GAM), 342
Generalized estimating equations (GEEs), 343
Generalized least squares (GLS), 344
Generalized linear mixed models (GLMM), 342
Generalized linear models (GLMs), 341
Gentoo penguins, 171
Geometric spreading, 171, 193
Geophony, 112, 218
Geophysical, 112
Giant pandas, 486
Gibbon, 98, 274, 466
Gleaning bats, 486
Gliders, 25
Global positioning system (GPS), 48
Goldfinches, 88
Goldfish, 365
Go/no-go response, 367
Goodness-of-fit tests, 345
Gorillas, 222
Graded calls, 405
Gramophone, 4
Grasshopper, 171, 466, 472
Gray treefrogs, 172, 475
Gray whale, 462, 465
Grazing angle, 195, 198
Greater bulldog bat, 426
Greater horseshoe bat, 174, 428, 430
Greater spear-nosed bat, 173
Great tits, 482, 483
Green sea turtle, 474
Green treefrog, 363, 382
Ground effect, 164, 165
Ground squirrels, 98
Grouper, 294
Guinea pig, 178, 356, 360
Gulf corvina, 238
Gunshots, 227

**H**
Habituation, 465, 473
Hair cell, 467, 487
Hankel function, 206
Harbor porpoise, 248, 377, 421, 429, 433, 436, 438, 439, 448, 465
Harbor seals, 98
Harmonic, 114, 406
Harp seals, 459
Hawaiian monk seal, 367, 369
Hearing damage, 478
Hearing loss, 176, 461, 485
Hearing threshold, 356
Heat capacity, 190
Helmholtz equation, 200, 204–206
Herons, 466
Herring, 477
Heterodyning, 60
Hidden Markov model (HMM), 298
Hierarchical modeling, 344
High duty-cycle bats, 428
High-pass filter, 55, 131
Hit rates, 366
Homogeneity of variances, 341
Honeybees, 220
Horse, 287
Horseshoe bat, 382, 486
House mouse, 356
House wrens, 484
Hoverflies, 220
Howler monkeys, 168
Hummingbird, 284
Humpback whale, 98, 105, 230, 231, 245, 271, 276, 304, 459, 489
Hurdle models, 343
Huygens' principle, 157, 161, 162, 167
Hydrophone, 21, 116
Hydrostatic pressure, 116
Hypercardioid, 44

**I**
Icebreaker noise, 188
Identifiers, 393
Image source, 197
Incoherent, 124
Incus, 484
Independent variables, 326
Indian elephant, 379, 380
Indian false vampire bat, 429
Indigo bunting, 96
Inferential studies, 321
Inferior colliculus evoked potentials, 375
Inflection point, 115
Informational masking, 461
Information entropy, 279
Infrasonic, 221
Infrasonic microphones, 43
Infrasound, 41, 111, 485
Inner ear, 358
Insects, 160, 220, 236, 471
Institutional Animal Care and Use Committee (IACUC), 88
Instrumentation microphones, 44
Integrated Nested Laplace Approximation (INLA), 345
Intensity, 194
Interactions, 330
Interference pattern, 164
Invertebrates, 468
Isosensitivity curve, 372
Isothermal layer, 192

**J**
Jackhammer, 226
Jays, 486

**K**
Kangaroo rats, 374
Katydids, 291

Killer whale, 92, 98, 232, 292, 295, 297, 367, 466, 476, 489
King penguins, 173, 460
Kit foxes, 362
Koalas, 174

**L**
Laboratory mice, 362
Lamb, 224
Larvae, 468, 470
Larynx, 391
Laser accelerometers, 18
Laser-doppler vibrometer, 76
Laser interferometers, 18
Laser microphones, 18
Law of reflection, 161
Leafhoppers, 472
Leaky modes, 205
Least-squares estimation (LSE), 336
Leatherback sea turtles, 473
Leks, 409
Leopard frogs, 476
Leopard seal, 231, 302
Level quantity, 123
Liberal response bias, 371
Limiting ray, 167
Linear regression models, 341
Line sources, 159
Link function, 341
Little brown bats, 173
Lizards, 474
Lloyd's mirror effect, 14, 197
Lobster, 229, 468–470
Local extremum, 115
Local maximum, 115
Local minimum, 115
Locust, 473
Logbook, 93
Loggerhead sea turtles, 474
Logit function, 342
Log-link function, 342
Lombard effect, 172, 174, 401, 466, 472, 475, 484, 486, 491
Longitudinal wave, 113
Longitudinal studies, 331
Long short-term memory, 293
Long-term spectral averages, 243
Lossless compression, 139
Loudness, 117
Low duty-cycle bats, 427–428
Low-pass filter, 131, 160

**M**
Machine learning, 290, 293
Mackerel, 478
Magnetaphone, 6
Magnitude, 116
Malleus, 484
Manatees, 465
MANCOVA, 341
MANOVA, 341
Marine mammal, 229
Markov Chain Monte Carlo (MCMC), 345
Marmoset monkeys, 175
Marmosets, 222
Masked threshold, 356
Masking, 358, 461, 472, 475, 491
Matched filters, 279
Mating calls, 363
Mauthner cells, 361
Maximum frequency, 115
Maximum likelihood estimation (MLE), 336
Mean-square sound pressure, 117
Measurement microphones, 18, 44
Mechanical turbulence, 170
Mechanical wave forms, 395
Mel-frequency cepstrum, 284
Melon, 434
Metadata, 88
Method of constant stimuli, 368, 381
Method of limits, 368, 369, 381
Mice, 274, 361, 373, 378, 487
Micro-electrical-mechanical system (MEMS), 43
Microphone, 116
Microphone array, 17
Microphonic potentials, 357
Middle ear, 358
Miners, 484
MiniDisc, 10
Minimum frequency, 115
Mink, 176
Missed detection, 280, 370
Mitigation, 463
Mixed layer, 190, 192
Mockingbirds, 106
Mode shape, 203
Modified method of limits, 369
Molecular relaxation, 160, 194
Molluscs, 467, 468
Mongolian gerbil, 356, 363, 377
Mongoose, 173, 467
Monophonic, 17
Mono recordings, 8
Moth, 422
Motifs, 274
Mountain sheep, 485, 486
Multi-collinearity, 345
Multilevel modeling, 344
Multiple regression, 341
Multivariate, 332
M-weighting, 130
Myotis bats, 286, 289
Mysticete, 229, 488

**N**
Narrowband, 131
Naval sonar, 186
Near-field, 126, 127

Needle electrodes, 377
Negative binomial distribution, 327
Negative phonotaxis, 363
Negative reinforcement, 366
Neotropical frog, 284
Nested variables, 344
Neuromasts, 477
Neutral response bias, 371
New River tree frog, 274
New Zealand fur seal, 223
Nightingales, 484
Nocturnal communication distances, 168
Noise, 111, 459
Noise control, 153
Nonlinear dimensionality reduction, 290
Non-parametric tests, 340
Non-song sounds, 276
Normal distribution, 327
Normalized difference soundscape index (NDSI), 248
Normal mode model, 203–205, 207, 235
Null hypothesis, 337
Numerical variables, 326
Nyquist frequency, 38, 137

**O**
Ocean acidification, 239
Octave band levels, 239
Octave bands, 132
Octopus, 468
Odontocete, 229, 488
Oilbird, 173, 430, 441, 442
Old-World monkeys, 487
Omnidirectional, 157, 171, 175
One-choice experiment, 364
One-tailed test, 337
One-third octave bands, 132
Open-reel recorders, 6
Operant conditioning, 366
Optical microphones, 44
Orange-eyed treefrog, 364
Organ of Corti, 461, 487
Orthopterans, 160
Ortolan buntings, 298
Oscillograms, 27
Oscilloscope, 19, 27
Otoacoustic emissions, 358, 373
Otoliths, 477
Ototoxins, 488
Outdoor sound propagation, 156, 157
Outer hair cells, 358
Outlier, 325
Oval squid, 377
Overdispersed, 342
Overtones, 114
Oysters, 460

**P**
Pacific white-sided dolphin, 102–103, 284
Pale spear-nosed bat, 223
Pallid bats, 173
Parabolic equation, 206, 207
Parabolic reflector, 5, 17
Parameters, 326
Parametric cluster analysis, 284
Parametric tests, 340
Paraxial approximation, 206
Particle acceleration, 120, 143
Particle displacement, 120, 143
Particle motion, 79, 361, 477
Particle velocity, 117, 120, 143
Passerines, 175
Passive acoustic monitoring, 22, 63
Passive bioacoustic studies, 37
Passive sonar, 21
Peak frequency, 131
Peak sound pressure, 117
Peak sound pressure level, 117
Peak-to-peak sound pressure, 116
Peak-to-peak sound pressure level, 117
Pearson's correlation coefficient, 333
Penguins, 482
Permanent threshold shift, 462
Phantom powering, 11
Phase, 113
Phonautograph, 3
Phonic lips, 434, 438
Phonometer, 44
Phonotaxis, 472, 475
Phons, 129
Photosynthesis, 238
Phrases, 274
Pickersgill's reed frog, 225
Piezoelectric, 51
Piezoelectric microphone, 16
Piglets, 486
Pigtail macaque monkey, 383
Pikas, 173
Pile driving, 234
Pilot whale, 232, 283, 290, 295, 489
Pine bark beetles, 472
Pinna, 358
Pinnipeds, 488
Pipistrelle bats, 174
Pitch, 113
Plague, 409
Plankton, 470
Plug-in-power (PIP), 11, 42
Plumbeous vireos, 483
Point estimate, 329
Point source, 127, 158, 159
Poisson distribution, 327
Polar bears, 223
Polar ice, 233
Polynomial, 342
Population consequences of acoustic disturbance (PCAD), 463
Population consequences of disturbance (PCoD), 463
Porpoises, 490

Positive phonotaxis, 363
Positive reinforcement, 366
Posterior distribution, 336
Power analysis, 339
Power quantity, 123
Power spectral density (PSD), 244
Power spectral density (PSD) percentiles, 141
Prairie dogs, 98, 292, 485
Prawn, 468
Preamplifier, 51
Precision, 324
Precision-recall, 282
Predation risk, 466
Predictive studies, 321
Predictors, 326
Prepulse inhibition, 362
Prepulse stimulus, 362
Presbycusis, 176, 356
Prescriber, 393
Pressure-release boundary, 197
Preyer reflex, 360
Primates, 222, 238, 361
Principal component analysis, 286, 334
Principle of parsimony, 345
Prior distribution, 336
Probability density, 142
Probability density function, 327
Probability distribution, 327
Probability mass function, 327
Production of vibrational signals, 396
Projector, 51
Propagation loss, 159, 186, 187, 193–195, 198, 199
Propagation modeling, 194
Psychometric function, 368
Psychophysical tuning curves, 383
Pulse length, 119
Pure tone, 114
Pygmy blue whales, 237, 245

**Q**
Quail, 484

**R**
Radiated noise level, 126
Radio-frequency microphone, 42
Railway, 225
Rain noise, 224
Random effects, 342
Random error, 323
Random forest, 294
Random variable, 326, 327
Range-dependent, 201
Range-independent, 201
Range of hearing, 358
Rarefactions, 113
Rat, 356, 361
Rattlesnakes, 92
Ray, 157, 201
Rayleigh roughness parameter, 197
Ray models, 201
Ray propagation, 157
Ray traces, 168, 169
Ray tracing, 157, 158, 207
Realizations, 327
Real-time spectrogram, 29
Received level, 159, 186
Receiver operating characteristic, 280, 370
Reception of mechanical signals, 396
Recurrent neural networks, 293
Red deer, 174, 298
Redundancy, 391
Reed warbler, 480
Reel-to-reel recorders, 6
Reflection, 161, 171, 187, 201, 235
Reflexive responses, 360
Reflex modifications, 362
Refracted waves, 196
Refraction, 166, 171, 187, 201, 235
Regression coefficient, 330
Regression models, 330
Reinforcement regimen, 371
Relative humidity, 160
Repeated measures study, 331
Research permits, 88
Residuals, 336
Response variables, 326
Reverberation, 103, 162, 300
Rhacophorid treefrogs, 475
Rhinoceros, 274
Right whale, 98, 231, 294
Risso's dolphins, 290, 295
Ritualization, 391
Rms bandwidth, 131
Road noise, 165
Rock hyraxes, 274
Rodents, 485
Root-mean-square sound pressure, 117
Ross seal, 231
Roughness, 199
Ruffed grouse, 161
Rufous-sided towhee, 93

**S**
Salinity, 192
Salinity profile, 192
Salmon, 478
Sampling frequency, 38, 137
Sampling rate, 3, 38, 137
Sampling units, 323
Savannah sparrows, 105
Scallop, 229, 469
Scattering, 161, 171, 187
Scattering loss, 197
Sciurids, 93
Scops owls, 174
Sea ice, 192
Sea lions, 223, 232
Seals, 223, 232

Seal scarers, 465
Sea otters, 223
Sea turtles, 474
Sea urchins, 229
Seismic airgun, 234, 302
Seismic communication, 396
Seismic survey, 233, 238, 470
Self-noise, 40
Shadow zone, 167, 169, 201, 227, 235
Shallow-water duct, 201
Shannon entropy, 279
Shannon-Nyquist sampling theorem, 3
Shear, 113
Sheep, 224
Ship noise, 240, 245
Shotgun microphones, 17
Shrew, 445
Shrimp, 228, 229, 239, 470
SI base units, 112
Sidelobes, 134, 140
Signal, 112, 393
Signal-to-noise ratio, 171, 186, 459
Significance level, 339
Sinusoidal, 117
SI system, 112
Smartphone, 32
Smooth functions, 342
Snakes, 220
Snapper, 478
Snapping shrimp, 240, 459
Snell's law, 166, 196, 201, 227, 235
Snowmobiles, 226
Social adaptation, 402
SOFAR channel, 202, 233, 235
Sonar equation, 186, 187, 208
Song, 219, 231, 274, 407
Songbirds, 274, 297
Sonobuoy, 54, 101
Sonogram, 27
Sonoran pronghorn, 485
Sound, 111, 113
Sound-attenuating chamber, 360
Sound detection, 363
Sound discrimination, 363
Sound exposure, 118
Sound exposure level, 118
Sound field, 158
Sound level meter, 44
Sound localization, 366
Sound maps, 170
Sound pressure, 116
Sound pressure level, 117, 159, 239
Sound propagation, 156, 227, 239, 245
Soundscape, 112, 141, 217, 218
Sound speed, 192, 199
Sound speed profile, 101, 166, 168, 192
Source level, 126, 159, 186, 187
Source-path-receiver model, 153
Southern brown treefrog, 475
Southern right whale, 273
Sparrows, 96
Spatial release from masking, 172, 178, 189, 401, 476, 483, 491
Species evenness, 246
Species-recognition, 96
Species richness, 246
Spectral density, 133
Spectral leakage, 140
Spectral probability density (SPD), 244
Spectrogram, 27, 136, 240, 241
Spectrogram cross-correlation, 278
Spectrogram equalization, 278
Spectrum, 131, 134, 136
Specular reflection, 161
Speed of sound, 120, 165
Spermaceti organ, 434
Sperm whale, 300, 435, 436
Spherical spreading, 159, 193, 195
Spline, 342
Spontaneous otoacoustic emissions, 373
Spotted dolphins, 297
Sprat, 478
Spreading loss, 187
Spring peepers, 236
Squid, 468, 469
Squirrelfish, 376
Squirrels, 238, 271, 466
Standard atmospheric pressure, 160
Stapedial reflex, 487
Stapes, 484
Starling, 178, 482
Start frequency, 115
Startle response, 477
Static information, 403
Statistical controls, 323
Statistical decision theory, 347
Statistical inference, 323
Statistical population, 323
Statistical power, 322
Statocyst, 467–470
Statolith, 467, 468
Stereocilia, 487
Stereophonic, 17
Stereo recording, 8
Stickleback, 362
Stonefly, 422
Streaked tenrecs, 445
Stress, 461, 467, 486, 491
Structure-borne, 112
Substrate-borne, 112
Substrate-borne vibrations, 75
Suction-cup electrodes, 377
Support vector machines, 284
Surface duct, 192, 201, 235
Survival models, 343
Swallow, 481
Swiftlet, 430, 442
Syllables, 274
Symbolic, 411
Synapses, 487
Synaptopathy, 487
Synergistic, 463

Systematic error, 323

**T**
Tail-flip reflex, 361
Tail-to-signal ratio, 162
Target motion analysis, 149
Target strength, 189
Tawny owls, 173, 224
t-distributed stochastic neighbor embedding, 290
Teager–Kaiser energy operator, 280
Technophony, 218
Telegraphone, 4
Teleost fish, 361
Temperature, 192
Temperature inversion, 168, 227
Temperature lapse, 167
Temperature profile, 168, 169, 192
Temporary threshold shift, 462
Tenrec, 445
Terminal buzz, 420, 423
Terrestrial biophony, 219
Terrestrial geophony, 225
Test statistics, 337
Theoretical model, 329
Thermal turbulence, 171
Thermocline, 55, 190, 192
Thunderstorm, 224
Timberline wrens, 275
Time-bandwidth product, 187
Time difference of arrival (TDOA), 144
Time domain, 134, 136
Time domain models, 200
Time series, 136
Torrent frogs, 475
Tortoises, 220, 473
Total internal reflection, 196
Towed array, 22, 54
Transient, 119
Transient otoacoustic emissions, 373
Transmission loss, 159
Transverse, 113
Treefrog, 176, 178, 475
Tree mice, 446
Tree-shrews, 238
Trucks, 225
True positives, 280
Truncated regression, 343
Tubeworms, 460
Túngara frogs, 174, 475
Turtles, 473
Two-choice experiment, 364
Two-tailed test, 337
Tympanal hearing, 471
Type I error, 339
Type II error, 339
Types of communication, vibrational signaler, 398

**U**
Ultrasonic, 220, 222
Ultrasonic microphones, 17, 43
Ultrasound, 41, 43, 111, 485
Uncertainty principle, 140
Unconditioned response, 357, 359, 365
Unconditioned stimulus, 365
Unidirectional, 44
Uniform manifold approximation and projection, 290
Univariate analyses, 332
Unmanned surface vessels (USVs), 25
Unmasked threshold, 356
Up/down staircase method, 368–370, 381–382
Upsweep, 115
Upward refracting, 166, 168
Upwind propagation, 167

**V**
Variables, 326
Vector sensors, 55
Velvet ants, 220
Ventriloquial calls, 406
Vespertilionid, 406
VHS tape recorders, 9
Vibrational behavior, 395
Vibrometers, 18
Vibroscape, 398
Video tape recorder, 9
Vireo, 88
Virtual source, 197
Viscous friction, 165
Vocalization, 396
Volcanic eruptions, 233
Volumetric array, 55
Voucher specimens, 106
48-V phantom power, 42

**W**
Walruses, 223, 232
Warbler, 294
Wasps, 220
Water-borne, 112
Wave equation, 156, 200
Waveform, 134
Wavefronts, 157
Waveguide, 201, 235
Wavelength, 114, 156, 161, 197, 227
Wavenumber, 205
Wavenumber integration, 205, 207
Wave rays, 157
Wax-cylinder recorder, 3
Weber fraction, 380
Weddell seal, 92, 98, 174
Weighting function, 488
Whale choruses, 243
Whale-watching boat, 186
Whistles, 231
White-crowned sparrows, 481, 484
White-footed mice, 465
White noise, 67
White-throated sparrow, 271

Wind noise, 224, 228, 232, 240
Windsock, 48
Wind velocity profile, 168, 169
Wireless microphones, 44
Wolves, 302, 486
Wood frogs, 476
Woodland caribou, 486

**Z**
Zebra finches, 297, 460, 481, 482, 484
Zebrafish, 362
Zero-crossing detection, 19
Zero-crossing points, 60
Zero-inflated models, 343
Zero-padding, 141
Zero-to-peak sound pressure, 117
Zero-to-peak sound pressure level, 117
Zooplankton, 470
Zoosemiotics, 389
Z-weighting, 129