Animal Signals and Communication 5

Mark A. Bee
Cory T. Miller *Editors*

Psychological Mechanisms in Animal Communication

Springer

Animal Signals and Communication

Volume 5

More information about this series at http://www.springer.com/series/8824

Mark A. Bee • Cory T. Miller
Editors

Psychological Mechanisms in Animal Communication

Springer

Editors
Mark A. Bee
Department of Ecology, Evolution, and Behavior
University of Minnesota - Twin Cities
Saint Paul, Minnesota
USA

Cory T. Miller
Department of Psychology, Neurosciences Graduate Program
University of California
La Jolla, California
USA

ISSN 2197-7305 ISSN 2197-7313 (electronic)
Animal Signals and Communication
ISBN 978-3-319-48688-8 ISBN 978-3-319-48690-1 (eBook)
DOI 10.1007/978-3-319-48690-1

Library of Congress Control Number: 2017931073

Printed on acid-free paper

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

Preface

Each of the two editors of this volume on *Psychological Mechanisms in Animal Communication* often sees the world of animal signals and communication very differently from the other editor. This is because each of us tends to view these topics through the eyes of our respective study organisms and the lenses of our past training. One of us (MAB), with a background in behavioral ecology, comparative psychology, and neuroethology, primarily studies acoustic communication in frogs. The other (CTM), with a background in evolutionary biology, cognitive psychology, and systems neuroscience, studies acoustic communication in monkeys. A number of years ago, however, at the First International Conference on Acoustic Communication by Animals, held at the University of Maryland in July 2003, we discovered that we share in common a deeply held opinion about research on animal communication: psychological mechanisms for signaling and receiving often get eclipsed by a fascination with signals.

Animal behaviorists, the media, and the public are frequently captivated by the elaborate and showy signals animals produce to communicate with each other. How can one fail to marvel at the music-like vocalizations of songbirds and humpback whales, the cacophony generated by choruses of calling frogs and insects, the dynamic and mesmerizing visual displays of a cuttlefish, the use of chemical signals by foraging ants, or the multisensory displays of peacock spiders and birds of paradise? Likewise, biologists and laypersons alike are often amazed by the diversity of functions that communication signals serve in the lives of animals, from courting and choosing mates to defending resources, strengthening social bonds, warning of danger, and begging for food or leading others to it. By comparison, the no-less elaborate and complex psychological mechanisms that have evolved for the purposes of signaling and receiving seem to inspire less wonder. One notable exception, of course, is the hunt for parallels and precursors of human language in other animals, which has been extensive and is ongoing. But many of the sensory, perceptual, and cognitive mechanisms that enable human communication are not language specific and can be studied in many other communicating animals. We hope this volume will stimulate wonderment regarding the diversity of psychological mechanisms for signaling and receiving found across a diversity of species.

After first crossing paths as graduate students at that 2003 conference, we set for ourselves the task of putting to pen our shared view on the state of research on psychological mechanisms in animal communication. It was only after several false starts, and after navigating the academic job market, that we finally accomplished our task in the form of an essay published in *Animal Behaviour* (Miller and Bee 2012). In that paper, we questioned whether the receiver psychology paradigm (Guilford and Dawkins 1991, 1993; Rowe 2013), which has been influential in shaping the field's thinking about the evolution of signal design and signaling behavior, has led to any useful, new insights into the psychology of receivers. Our general conclusion was that it had not. We argued that the receiver psychology paradigm could have a much greater impact on the study of animal signals and communication if it also promoted research on the actual psychological mechanisms involved in communication and their potential evolutionary diversity. Specifically, we advocated for seeking a broader and deeper understanding of receivers' "psychological landscapes" through the comparative study of the sensory, perceptual, and cognitive mechanisms that enable receivers to acquire, process, and act upon signals. This volume represents an extension of the basic thesis of our 2012 essay: understanding animal signals and communication depends fundamentally on understanding the psychological mechanisms operating in signalers and receivers.

In putting this book together, we have had the fortunate pleasure to engage as contributing authors a number of our colleagues whose research programs on animal communication have focused on various aspects of sensation, perception, and cognition. Collectively, their chapters span an impressive range of topics, taxa, and tools. We are deeply grateful to each author for their contribution to this volume and, more broadly, for their contribution to making new discoveries about signaler and receiver psychology. Their research inspires our own. We are also grateful to the series editors, Vincent Janik and Peter McGregor, for inviting us to edit this volume and for their input and guidance both on the large, big-picture issues involved with putting together an edited volume and on the small, mundane issues that are an inevitable part of doing so. We thank Andrea Schlitzberger, from Springer, for her help and dedication to coordinating all aspects of this project. The National Institutes of Health and the National Science Foundation have generously supported our research on animal communication over the years. Finally, we thank our families—Meggan, Karis, Abigail, River, and Sierra—for their enduring support.

Mark A. Bee
St. Paul, MN, USA

Cory T. Miller
La Jolla, CA, USA

References

Guilford T, Dawkins MS (1991) Receiver psychology and the evolution of animal signals. Anim Behav 42:1–14

Guilford T, Dawkins MS (1993) Receiver psychology and the design of animal signals. Trends Neurosci 16(11):430–436

Miller CT, Bee MA (2012) Receiver psychology turns 20: is it time for a broader approach? Anim Behav 83(2):331–343

Rowe C (2013) Receiver psychology: a receiver's perspective. Anim Behav 85(3):517–523

Contents

Chapter 1
Signaler and Receiver Psychology

Mark A. Bee and Cory T. Miller

Abstract This edited volume on *Psychological Mechanisms in Animal Communication* highlights research on the sensory, perceptual, and cognitive mechanisms that underlie signaling and receiving. It brings together researchers working on a broad range of conceptual questions in diverse animal systems and using an assortment of empirical tools. Collectively, these researchers seek to understand how signalers signal and receivers receive. This introductory chapter introduces the major questions in studies of signaler and receiver psychology that are explored in greater depth in subsequent chapters. In so doing, this chapter makes the case that a research agenda aimed at elucidating the mechanisms of signaler and receiver psychology complements and enriches several current areas of animal communication research, in particular those focused on signal design and the parallels and precursors of human language in animals. Ultimately, the goal of this volume is to lay a solid foundation for broader and more comparative studies that investigate the psychological mechanisms of animal communication.

1.1 Sights and Sounds of Spring

Across a large swath of Canada and the United States, the arrival of migrant red-winged blackbirds, *Agelaius phoeniceus* (Fig. 1.1), at the close of winter, is among the earliest and surest signs of spring. Perhaps the most common and best-studied songbird in North America (Orians and Christman 1968; Yasukawa and Searcy 1995; Beletsky 1996; Beletsky and Orians 1996), red-winged blackbirds are

---

M.A. Bee (✉)
Department of Ecology, Evolution, and Behavior, University of Minnesota – Twin Cities, St. Paul, MN, USA
e-mail: mbee@umn.edu

C.T. Miller
Department of Psychology, Neurosciences Graduate Program, University of California, San Diego, La Jolla, CA, USA
e-mail: corymiller@ucsd.edu

M.A. Bee, C.T. Miller (eds.), *Psychological Mechanisms in Animal Communication*, Animal Signals and Communication 5,
DOI 10.1007/978-3-319-48690-1_1

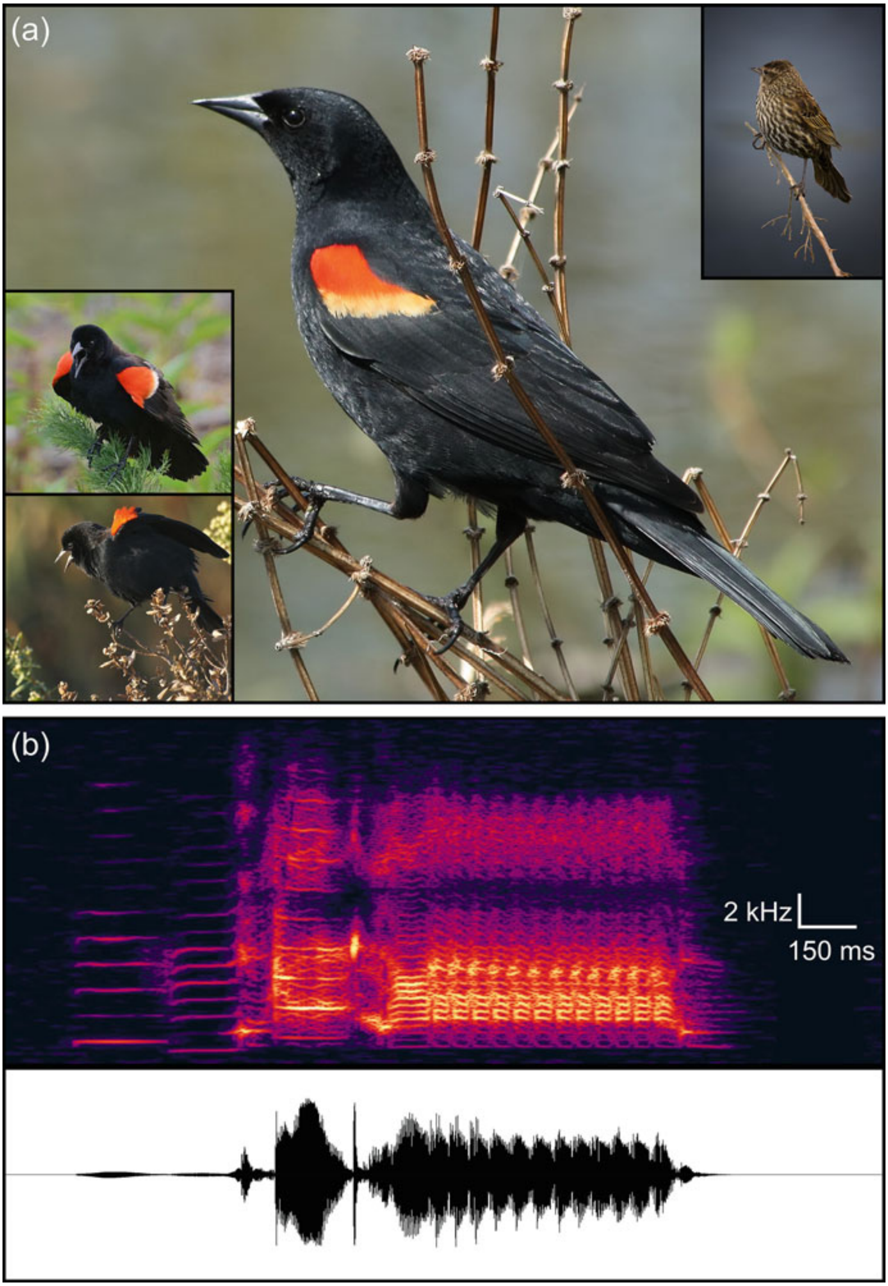


Fig. 1.1 Visual and acoustic signaling in the red-winged blackbird, *Agelaius phoeniceus*. (**a**) Adult male showing the red and yellow wing epaulet and solid black body. *Insets* in the *lower left* illustrate multimodal signaling in action in the form of vocally signaling males engaged in simultaneous visual displays of the wing epaulet. *Inset* in the *upper right* shows a cryptically colored female. (**b**) Spectrogram (*top*) and waveform (*bottom*) of a typical red-winged blackbird song (*oak-a-lee*). Males often sing this song while performing the "song spread" visual display illustrated by the *top left inset* in (**a**). *Attributions*: The large figure in (**a**) courtesy of Walter

sexually dimorphic and highly polygamous. Males are dark black with a boldly colored wing epaulet, and females are mottled brown and more cryptically colored (Fig. 1.1a). Males defend territories in wetland and upland habitats, where they can often be observed proudly perched atop vegetation. Chases and sometimes fights occur as males establish mutually respected territory boundaries with their neighbors. Up to 15 females may nest in a single male's territory and mate with him, but they also solicit copulations from neighboring males. Although the species' common name is derived from the prominent red and yellow wing epaulet borne by adult males, the epaulet is just one of the many sights and sounds in the species' impressive repertoire of signals.

As songbirds, male red-winged blackbirds learn to produce their acoustically complex songs (*oak-a-lee*; Fig. 1.1b). The song is both species-specific and individually distinct, and it functions in both territory defense and mate attraction. Both sexes respond more strongly to red-winged blackbird songs than to the songs of other species, and they also discriminate between local and foreign regional dialects. Territorial males learn to recognize nearby neighbors by the individually distinctive features of their songs. Females also sing, either in response to males, to announce their arrival and departure from a male's territory, or during aggressive interactions with other females. "Song," however, is just one of several acoustic signals produced by these voluble birds. "Courtship calls" are used preceding mating, "contact calls" are produced while foraging, "feeding calls" are given when returning to the nest to feed offspring, "flight calls" are uttered by males when they leave their territory, "threat calls" are directed toward opponents in aggressive contests, and "alarm calls" are produced in response to predators. The species reportedly produces seven different types of alarm calls, some variants of which may be specific to certain types of predators, such as hawks.

As a visual signal, the eponymous wing epaulet contrasts with the solid black of the male's body and the greens and browns of the surrounding habitat (Fig. 1.1a), but it is not a static visual signal. Rather, males can accentuate the epaulet by erecting its feathers or conceal it by covering it with black feathers. This behavioral flexibility makes the epaulet one of several dynamic visual signals in the species' rich signal repertoire. The epaulet is perhaps most obviously on display when males perform the conspicuous "song spread" display (Fig. 1.1a, *top left inset*), in which

Fig. 1.1 (continued) Siegmund (licensed under the Creative Commons Attribution-Share Alike 3.0 Unported, 2.5 Generic, 2.0 Generic and 1.0 Generic license; commons.wikimedia.org/wiki/File:Agelaius_phoeniceus_0110_taxo.jpg). The *top inset* at the *lower left* courtesy of Geoff Gallice (licensed under the Creative Commons Attribution 2.0 Generic license; commons.wikimedia.org/wiki/File:Flickr_-_ggallice_-_Red-winged_blackbird.jpg); the *bottom inset* at the *lower left* courtesy of Ingrid Taylar (licensed under the Creative Commons Attribution 2.0 Generic license; commons.wikimedia.org/wiki/File:Blackbird_Calling.jpg). The *inset* at the *upper right* courtesy of Steven Pavlov (licensed under the Creative Commons Attribution-Share Alike 3.0 Unported license; commons.wikimedia.org/wiki/File:Agelaius_phoeniceus,_female.jpg). The spectrogram and waveform in (**b**) were created using a song recording provided courtesy of Keith Barker at the University of Minnesota the male sings while simultaneously erecting these feathers, lowering the tail, and extending the wings and spreading them slightly apart from the body. Males can modulate the intensity of their song spread displays depending on the behavioral context. They can also expose or exaggerate the epaulet and direct it toward either females or rival males using a number of other visual displays, such as the “bill-up display,” the “defensive flutter,” or a “crouch” posture, as well as various flight displays, such as “flight song” and “fluttering flight.”

Through decades of research on red-winged blackbirds, we have learned a great deal about its acoustic and visual signals and their social and sexual functions (reviewed in Yasukawa and Searcy 1995). Indeed, we now know a great deal about signals and their functions in a great many animal species (Bradbury and Vehrencamp 2011). Various fields devoted to studying animal signals and communication, such as behavioral ecology and neuroethology, have made extraordinary advances in understanding signal design and the mechanisms by which signals are produced and transduced. A particularly rich and productive focal area of research in behavioral ecology has been the identification of diverse sources of selection acting to shape signal design (Hebets and Papaj 2005; Hebets et al. 2016). Yet, a host of unanswered questions remain regarding the psychological mechanisms underlying communication in animals like the red-winged blackbird. These questions pertain less to signal design and more to signal usage (by signalers) and signal processing (by receivers). Signaling and responding to signals are behavioral actions that engage a suite of evolved psychological mechanisms. The dual purpose of this volume is to highlight current research, and stimulate future research, on these mechanisms.

1.2 Psychological Mechanisms in Animal Communication

In this volume, we use the phrase *psychological mechanisms* to refer broadly and collectively to all of the processes carried out by the neural and neuroendocrine mechanisms operating in the peripheral and central nervous systems that are responsible for transducing, coding, processing, decoding, selecting, storing, retrieving, comparing, and acting upon information in signals. The phrase is meant to encompass the entire breadth of sensory, perceptual, and cognitive processes that underlie a signaler’s abilities to adaptively use the signals in its repertoire and a receiver’s abilities to adaptively respond to them. This inclusive view of psychological mechanisms in animal communication spans a broad array of interrelated phenomena, such as sensory processing, perceptual object formation, categorization, social cognition, numerical cognition, learning, memory, attention, decision-making, and concepts, among others. With this volume, we intend to shine a light on research programs that reach beyond questions of signal design and function to elucidate sensory, perceptual, and cognitive processes involved in signaling and receiving across a wide diversity of animals. We return to red-winged blackbirds to introduce the specific topics covered in subsequent chapters.

1.2.1 Variation in Sensory Processing

In red-winged blackbirds, as in other species, communication signals vary across many different scales. Both the song and plumage of red-winged blackbirds are species-specific and can be used as diagnostic cues for species identification. Their acoustic and visual signals also differ between the sexes, as well as among individuals of the same sex, which allows territorial males, for example, to discriminate between neighbors and strangers. Within individuals, signals and signal usage vary seasonally. Male red-winged blackbirds, for example, sing more and bear more striking plumage during the breeding period than at other times of the year. The important roles in animal communication of signal variation across these different scales—species, sexes, seasons, and individuals—are well established. But what about variation across these same scales in the sensory mechanisms that transduce and code signal information? This important question, which has received considerably less attention in the literature than corresponding work on signal variation, is taken up by Henry et al. in Chap. 2.

Henry et al. (Chap. 2) review comparative studies of songbirds that have used minimally invasive neurophysiological tools to examine species, seasonal, sex, and individual differences in sensory mechanisms. These studies have uncovered a high degree of variation in auditory processing that is largely consistent with the "sender–receiver matching hypothesis," which holds that coevolution of signals and sensory systems should result in a good match between signal structure and the tuning of relevant sensory systems. In the case of songbirds, this match exists between the spectral and temporal features of song and the neurosensory coding strategies used to process these features. This body of work has uncovered species, seasonal, and sex differences that potentially enhance perception of communication signals in ways shaped by a species' ecological requirements and seasonal and sex-specific use of signals. More recent work is uncovering surprising levels of individual variation in signal processing strategies too, the broad implications of which are still coming into focus. The studies reviewed by Henry et al. (Chap. 2) dispel once and for all any notion that receivers represent a uniform and invariant group of individuals that process signals the same way. Just like signals, sensory processing strategies vary across multiple scales. Multi-scale variation in signal processing has profound implications for understanding animal communication.

1.2.2 Signals and Signalers as Perceptual Objects

As we have argued elsewhere (Miller and Cohen 2010; Miller and Bee 2012), studying the psychological mechanisms of animal communication requires that we adopt a new view of signals and signalers as perceptual "objects" (Spelke et al. 1993; Kubovy and Van Valkenburg 2001; Scholl 2001; Van Valkenburg and Kubovy 2004). As the basic unit of perception, objects are formed by binding stimulus features (x, y, and z) into a coherent representation that can be segregated from other potential objects in the environment. Object perception arises from neural mechanisms that bind together these separate features (e.g., color, shape, size). Notably, object formation is not the same as detecting the presence of an object in the environment or discriminating between Object 1 and Object 2 based on differences in their features (x_1, y_1, z_1 vs. x_2, y_2, z_2). In the literature on human perception, particularly visual processing, objects are considered to be spatiotemporally bounded feature clusters (Scholl 2001). Physical entities present in visual scenes can be perceived as "visual objects," for example, when they move through space and time as a single coherent, bounded unit. Sounds present in acoustic scenes can also be perceived as "auditory objects" (Griffiths and Warren 2004; Bizley and Cohen 2013). Many of the same or related Gestalt grouping principles that guide the formation of visual objects appear to function in forming auditory objects too (Bregman 1990; Dent and Bee in press). Within this context of perceptual objects, it becomes critical to ask, how do receivers integrate complex mixtures of sensory inputs, potentially across multiple modalities, to form coherent perceptual objects of signals and signalers? We can reframe this question from the perspective of a female red-winged blackbird: how does she integrate her perception of a male's black coloration, his wing epaulet, his movements of wings and tail in a song spread display, and the different spectral and temporal elements of his song (Fig. 1.1) into a cross-modal perceptual object that corresponds to a potential mate?

The issue of signals and signalers as perceptual objects is taken up in two chapters. Klump (Chap. 3) reviews research on auditory scene analysis in another well-studied songbird, the European starling. With the starling communication system as a backdrop, this research has combined psychophysical and neurophysiological methods to identify the acoustic cues used to perceptually organize complex acoustic scenes and the neural mechanisms for coding these cues. The chapter focuses on how features of background noise, like that of a dawn chorus or nighttime roost, can be exploited to improve signal perception, and how temporally concurrent signals, such as overlapping songs, can be perceived as separate auditory objects. The key take-home messages from this chapter are that starlings and humans exploit many of the same perceptual cues to construct auditory objects and that starlings might serve as excellent models for understanding the neural basis of perceptual organization. Farris and Taylor (Chap. 4) continue the theme of signals and signalers as perceptual objects. They review work on perceptual grouping in the context of mate selection, primarily in anurans but also in orthopterans. The first part of their chapter explores the cues that frogs (and crickets) use to perceptually bind different structural elements of acoustic signals into coherent auditory objects. Farris and Taylor (Chap. 4) illustrate that, while many of the same cues are used by humans and other animals, there is also potentially important species diversity in perceptual grouping. In the second part of their chapter, Farris and Taylor (Chap. 4) turn to multicomponent signals and multisensory integration. They describe a research program that has used robotic frogs to investigate the principles of cross-modal grouping in the context of animal communication. Specifically, they detail behavioral studies that have attempted to elucidate the rules governing audiovisual integration in forming cross-modal objects that correspond to potential mates whose signals contain multiple components in different sensory modalities.

Research on the psychology of receiving multicomponent signals is fundamentally important because it complements other research agendas focused on the evolution of complex signals. Indeed, one lingering question about the evolution of multicomponent signals is that they often seem advantageous from a psychological perspective, but disadvantageous from an economic one that explicitly considers their benefits and costs. In Chap. 5, Rubi and Stephens develop an economic model that resolves this apparent paradox. Their model shows that under certain circumstances, such as when combinations of signal components indicate the occurrence of a rare event (e.g., a particularly high-quality mate), it can be economically, and not just psychologically, advantageous for receivers to attend to multicomponent signals. Together, the chapters by Klump, Farris and Taylor, and Rubi and Stephens highlight the value of taking comparative, psychophysical, neurophysiological, and theoretical approaches to investigate how and why receivers might be adapted to form perceptual objects of complex signals.

1.2.3 Categorization and Social Cognition

A category can be said to exist whenever two or more discriminable objects or events are treated equivalently (Mervis and Rosch 1981). Signaling and receiving by red-winged blackbirds demonstrate the importance of categories in communication. For example, these birds have a category for "local conspecific song," because they respond to songs from conspecific males from the same geographic region in much the same way, but in a way that is different from how they respond to the songs of conspecific males singing other regional dialects or males of other species. Territorial males treat variation in the different renditions of song from a neighbor as "neighbor song," a category that excludes the songs of unfamiliar males. If blackbirds signal using distinct alarm calls upon detecting a hawk, then they have a "hawk" category specific to this type of predator. Thus, correct categorization of objects and events is a fundamental cognitive challenge faced by both signalers and receivers. Social categorization is a critically important process that refers to the ability to assign conspecifics to different social categories, such as dominant versus subordinate, kin versus non-kin, and mate versus non-mate, based on patterns of individual variation in their phenotype, including signals. Categorization, generally, and social categorization, specifically, are topics covered by several chapters in this volume.

Akre and Johnsen (Chap. 6) carefully examine how proportional processing consistent with Weber's law can lead to receiver errors in signal categorization. Proportional processing refers to the idea that receivers categorize two or more signals as being the same or different based on their proportional, not their absolute, differences. A diversity of examples, from damselflies to deep-sea fish, are provided to illustrate how proportional processing works and how it can act as a source of selection shaping the evolution of signal design and the behavioral strategies of signalers, as well as the evolution of alternative receiver mechanisms. Bee (Chap. 7) critically reviews previous work on learned social categorization in frogs. Like red-winged blackbirds, some territorial frogs discriminate between the vocal signals of neighbors and strangers. Studies of North American bullfrogs indicate that merely hearing the calls of a neighbor coming repeatedly from a particular direction allows territory holders to encode enduring representations of the information necessary to categorize an individual as a neighbor. Social categorization is also a key topic addressed in chapters by Manser (Chap. 8) and Zuberbühler (Chap. 9). In her review of field studies of communication and social interactions in meerkats and banded mongooses, Manser (Chap. 8) describes how these highly social animals categorize conspecifics at multiple scales that include kinship, dominance rank, social group membership, and recognition of specific individuals. In his chapter, Zuberbühler (Chap. 9) discusses many of these and related issues of social categorization with respect to work on nonhuman primates. The treatments of social categorization by Manser (Chap. 8) and Zuberbühler (Chap. 9) occur within a somewhat broader discussion about referential signals, semantics, and concepts, three key topics we return to in the next section.

Important questions about social cognition emerge when the social category of interest is so narrow as to correspond to a single individual (Seyfarth and Cheney 2009, 2015a, b). This form of social categorization is usually referred to as "individual recognition." The exceptionally broad usage of that descriptor in the literature, however, belies our ignorance of the phenomenon's underlying mechanisms: when we say that bullfrogs, blackbirds, banded mongooses, and baboons all "recognize" other individuals, we surely do not wish to imply that precisely the same psychological mechanisms are at work and that these animals share the same perceptual and cognitive experiences. Upon demonstrating that some form of recognition exists, several interesting and empirical questions for further study arise (Bee 2006). What sensory, perceptual, and cognitive mechanisms are involved? Are representations of individuals multimodal, and is multimodality required for individual recognition? How enduring and specific are such representations? What additional elements of social knowledge about recognized individuals are also encoded (e.g., fighting ability, signal repertoire, dominance rank, social bonds, and familial relationships)? And how do answers to these various questions differ across species? Readers are encouraged to bear these sorts of questions in mind while reading chapters that deal with issues of individual recognition and social cognition (Chaps. 7–9).

1.2.4 Referential Signals, Semantics, and Concepts

If a red-winged blackbird sees a hawk and produces a distinctive, hawk-specific alarm call, to what extent can it be said to possess a mental representation of a hawk? And if a blackbird receiver hears a hawk-specific alarm call, to what extent does the call evoke a mental representation of a hawk in the receiver versus merely eliciting the most adaptive escape response? Does the blackbird signaler have *beliefs* about the receiver's mental representations of hawks, and does it *intend* to evoke a representation of a hawk? That is, does a hawk-specific alarm call function like a word in human language? Do blackbirds have a theory of mind? Questions about the extent to which signals refer to events or objects in the environment external to the signaler are central to ongoing debates about referential signaling, semantics, and concepts in animal communication (Newen and Bartels 2007; Wheeler and Fischer 2012, 2015; Stegmann 2013; Andrews 2015; Scarantino and Clay 2015; Seyfarth and Cheney 2015a). These issues have played key roles in research on the cognitive abilities of nonhuman primates and a few other mammals, and they are taken up by two chapters in this volume.

Manser (Chap. 8) tackles the issues of referential signaling and semantics by drawing on long-term studies of communication in meerkats and banded mongooses. She proposes an expanded framework for referential signals that considers three different types of referents that span a range of communication signals and contexts. Signals can refer to external objects or events (e.g., the presence of predators or food). Such signals are typically referred to as *functionally* referential as a way to remain agnostic about the degree of similarity between the cognitive mechanisms underlying their usage and perception and those underlying human language. According to Manser's framework, signals may also refer to the signaler's individual traits or social category membership. Signals that are age- or sex-specific that reflect a signaler's dominance rank or that function as identity signals represent examples of this type of referent. Finally, signals may refer to the current behavior of the signaler or its motivational and emotional states. The tonic use of contact calls in safe foraging contexts would represent an example of this type of referent. Manser (Chap. 8) explores potential similarities and differences in the mechanisms underlying the usage and perception of these different types of referents. Many of these same ideas permeate Zuberbühler's (Chap. 9) chapter on social concepts and communication in nonhuman primates. Drawing primarily on studies of chimpanzees, baboons, Diana monkeys, and Campbell's monkeys, he explores primates' concepts of group identity, dominance, and social bonds (including third-party relationships). Against this backdrop, he also reviews studies investigating how nonhuman primates use communication to recruit friends, to signal about food, and to warn others of threats from predators. Zuberbühler (Chap. 9) links concepts in nonhuman primates to the evolution of large brains (for a given body size) in primates and the conceptual foundations of human language. Together, these chapters on meerkats, mongooses, monkeys, and apes illustrate the potentially important roles of higher-level cognitive processing in the communication systems of diverse nonhuman animals.

1.2.5 Decision-Making

Decisions represent the outcomes of processes by which nervous systems translate sensation into behavioral actions (Pearson et al. 2014). Communicating animals face a bewildering array of decisions to make. Signalers must make decisions about whether, what, and when to signal; receivers must decide whether, how, and when to respond. For example, when another male begins singing near the border of a male blackbird's territory, how does the resident decide whether the song was produced by a neighbor versus a stranger and whether the song should be ignored or respond to aggressively? And if it is a stranger, can the resident decide to respond in a way that ensures his signals of aggression are appropriately directed toward the intruder and do not cause collateral damage to his relationships with females or adjacent neighbors? A male might also face potential trade-offs in deciding to signal his opposition to an intruder instead of signaling his acceptance of a nearby female, as the most appropriate signals for aggression and courtship usually differ. With the owner of her territory distracted, how does a female decide whether and when to signal to a neighboring male about her willingness to copulate with him?

The importance of decision-making as a cognitive process in animal communication is widely recognized and has received some attention in the literature, particularly in the context of mate choice (Ryan et al. 2009). Yet many open questions remain regarding how signalers and receivers make critical decisions both as communicators and as eavesdroppers. The overall theme of decision-making is intricately woven through most chapters of this volume, though it is more explicit in some than in others. Akre and Johnsen's (Chap. 6) discussion of proportional processing illustrates one important reason why receivers can make wrong decisions and how receiver errors can influence the evolution of signal design, signaler behavior, and receiver mechanisms. Bee (Chap. 7) reviews work on identifying receiver decision rules for discriminating between neighbors and strangers and how these decision rules relate to patterns of individual variation in signals. In Chap. 10, Toarmino et al. review work on the role of decision-making in the antiphonal calling behavior of common marmosets. Specifically, they examine the decisions of adults, and the ontogeny of decisions in juveniles, regarding whether to signal, what to signal, and when to signal within the complex ecological and social landscapes inhabited by this Neotropical, group-living monkey. Page and Jones (Chap. 11) examine decision-making from the perspective of a well-studied eavesdropper. They describe results from studies of the perceptual and cognitive mechanisms underlying a frog-eating bat's decision to attack prey items in the face of considerable uncertainty. Of particular interest in this chapter is research examining how receivers integrate multiple sources of information—passive listening, echolocation, chemosensation, and social learning—not only when making decisions about attacking potential prey but also in flexibly altering these decisions in the face of new information.

1.2.6 Learning and Memory

Learning and memory are key elements of psychology that play critical roles in animal communication. Frequently, the roles of learning and memory in communication are discussed in the context of vocal learning (Janik and Slater 1997; Doupe and Kuhl 1999; Soha and Peters 2015). Like other songbirds, for example, male red-winged blackbirds learn to sing their songs as juveniles by listening to, and memorizing, songs produced by adult males. Both the existence of regional song dialects and the ability of both males and females to discriminate local from foreign dialects further reflect the operation of learning and memory. That male blackbirds also discriminate between neighbors and strangers based on song indicates they learn about and remember the songs of other individuals and use this information in making strategic decisions about territory defense. In some songbirds, this learned familiarity with a neighbor's song is long-lasting and endures across multiple breeding seasons (Godard 1991).

While vocal learning is an important element of communication for some animals, it is also one well covered previously in the literature (Janik and Slater 1997; Doupe and Kuhl 1999; Soha and Peters 2015); therefore, it is not a subject covered explicitly in this volume. Instead, learning and memory in communication, apart from vocal learning, are discussed in a variety of different contexts in several chapters. For example, Klump (Chap. 3) discusses how familiarity gained from either learning to sing a song or hearing others sing it influences starlings' ability to perceptually restore masked portions of a song. Rubi and Stephens (Chap. 5) briefly discuss learning in the context of their economic model of multicomponent signaling. Bee (Chap. 7) describes work on habituation as a form of long-term, stimulus-specific, and context-specific learning that allows territorial frogs to direct lower levels of aggression toward neighbors in appropriate places. Learning and memory are implicit in Manser's (Chap. 8) discussion of referential signaling and semantics in meerkats and in Zuberbühler's (Chap. 9) treatment of concepts and communication in nonhuman primates. Toarmino et al. (Chap. 10) discuss recent work examining the experiential acquisition of turn-taking behavior in the antiphonal calling of marmosets, a learned behavior that appears to involve instructional feedback from parents. Page and Jones (Chap. 11) review recent work examining the role of social learning (i.e., learning by observing others) in the eavesdropping and foraging behaviors of frog-eating bats. Together, these chapters serve to demonstrate the vitally important and diverse roles of learning and memory in many animal communication systems.

1.3 Toward a Psychology of Signaling and Receiving

It is an exciting time for scientists interested in the psychological mechanisms of animal communication. Significant questions remain unanswered, and many research avenues still remain largely unexplored. One could argue, as indeed we have elsewhere (Miller and Bee 2012), that progress thus far has been limited, in part, because the field of animal communication, particularly research on non-primate species, tends to place what we consider to be disproportionate emphasis on matters of signal design and function (e.g., Maynard Smith and Harper 2003). These "signal-centered" views of communication reflect the large impact and relatively narrow focus (at least where causal and developmental mechanisms are concerned) of the field of behavioral ecology. However, even many neuroethological studies of communication emphasize mechanisms directly related to signal design, such as how signals with specific features are produced (e.g., mechanically and by central pattern generators) and how these features are transduced (e.g., by the peripheral nervous system) and coded (e.g., by feature detectors in the central nervous system). As a result, much of the current literature (and of course there are notable exceptions) reflects an unnecessarily narrow view of communication. Perhaps it is not surprising that external, showy signals attract the attention of intended receivers and scientists alike. But we believe greater effort directed toward elucidating the internal, psychological mechanisms involved in producing, using, and responding to signals complements signal-centered approaches by providing a more encompassing framework for investigating the full range of mechanisms that underlie animal communication.

Among the notable exceptions to signal-centered views are research agendas investigating communication and cognition in primates and vocal learning in songbirds, both of which have been motivated, in part, by efforts to understand the mechanisms and evolution of human language (Doupe and Kuhl 1999; Cheney and Seyfarth 2010; Fitch 2010; Shettleworth 2010; Soha and Peters 2015). While extremely productive and informative, these taxon-specific and language-related research programs often shine dim light at best on other psychological mechanisms in communication. On the surface, another notable exception would appear to be the influential receiver psychology paradigm (Guilford and Dawkins 1991, 1993; Rowe 2013). However, by reducing the rich "psychological landscape" of receivers to considerations of a signal's detectability, discriminability, and memorability, the receiver psychology paradigm is having a limited impact on advancing new knowledge about psychological mechanisms in animal communication. Instead, it poses as a lens through which psychological mechanisms are viewed as sources of selection acting on signal efficacy. We have previously advocated for a broader view of receiver psychology (Miller and Bee 2012), and our hope is that this volume will go some way toward stimulating the desired expansion of scope.

To advance research on animal communication, we need a comparative and integrative approach to signaling and receiving that examines a diversity of psychological processes, such as those outlined in Sect. 1.2, in a diversity of species and using a diversity of empirical methods. The goals of such an approach must be to reveal how sensory, perceptual, and cognitive mechanisms drive signal usage and receiver responses and to discover how and why these mechanisms differ across species in light of differences in the species' ecological and social landscapes and phylogenetic histories. Collectively, the chapters of this volume provide a roadmap for reaching these goals. Readers are encouraged to keep a watchful eye open for both explicit and implicit species comparisons. For example, Henry et al. (Chap. 2) and Bee (Chap. 7) place their reviews of species differences in phylogenetic frameworks. Klump (Chap. 3) and Farris and Taylor (Chap. 4) make explicit comparisons of auditory scene analysis between starlings and humans, and between frogs and humans, respectively. Manser (Chap. 8) compares referential signaling in meerkats and other animals, especially nonhuman primates, and Zuberbühler (Chap. 9) discusses humanlike concepts across several nonhuman primate species. Likewise, the chapters of this volume illustrate the effective use of diverse empirical tools for elucidating psychological mechanisms. Henry et al. (Chap. 2) review results from noninvasive electrophysiological recordings of scalp potentials. Klump (Chap. 3) describes a research agenda that integrates operant conditioning with multiunit and single-unit recordings of cortical neurons. Farris and Taylor (Chap. 4) and Page and Jones (Chap. 11) illustrate the largely untapped promise of robotics to investigate psychological mechanisms in animals. Rubi and Stephens (Chap. 5) show how theoretical modeling efforts can bridge studies of animal communication from psychological and behavioral ecological perspectives. Several chapters show how the workhorse of animal communication research—the trusty playback study—can be used to probe psychological mechanisms in field studies (Chaps. 6–9) and in highly controlled laboratory experiments (Chaps. 4, 6, 10, and 11). Toarmino et al. (Chap. 10) review the use of interactive playback experiments designed to create virtual animals in the laboratory. Implementing research programs that successfully deploy this broad array of tools across a diversity of species will demand future collaborations between animal behaviorists, psychologists, neurophysiologists, engineers, and roboticists.

1.4 Summary and Future Directions

We began this chapter with a signal-centered description of visual and acoustic communication in red-winged blackbirds. Against this backdrop, we introduced how the chapters in this volume collectively outline a field of study devoted to elucidating the diversity of sensory, perceptual, and cognitive mechanisms that are engaged when animals communicate. These mechanisms, like signals, have evolutionary histories and are shaped by natural and sexual selection pressures. Yet, in the study of animal communication, these mechanisms are rarely made explicit in signal-centered treatments, nor is their potential breadth made readily apparent by studies that seek parallels between human language and other forms of animal communication. The methods best used to study the psychological mechanisms of communication are simultaneously comparative and integrative and currently include field and laboratory experiments, interactive experiments, highly controlled psychophysical experiments based on conditioning, electrophysiological recordings of neural activity, and robotics.

We have asked authors to conclude their chapters with short summaries and outlines for future research directions. Readers are, therefore, referred to the more extensive and specific ideas for future research found at the end of each subsequent chapter to gain a sense of where this field is going. Here, we end with a call for greater research not only into the animal systems and psychological processes covered in this volume but also into systems and processes that may be noticeably absent to some readers. For example, readers interested in learning more about the psychology of communication in invertebrates, such as insects and cephalopods, as well as some vertebrates, such as fish and reptiles, will be disappointed in the extent of coverage of these animals in the following pages. The same holds for readers interested in chemical, electrical, and seismic communication. The chapters of this volume are admittedly biased toward acoustic signaling in songbirds, frogs, and mammals. While biased, this coverage accurately reflects much of the current state of research on the psychological mechanisms of animal communication. Admittedly, the bias also reflects the research interests and expertise of the volume editors, as well as the unavailability of some colleagues to contribute an invited chapter. At the same time, this volume does not cover exhaustively all of the psychological processes with critical functions in signaling and receiving. Diverse forms of numerical cognition, for example, are no doubt important in a broad range of species and communication contexts, such as call matching in frogs (Gerhardt et al. 2000) and group size assessment in primates (Wilson et al. 2001; Kitchen 2004), but this interesting topic is not covered explicitly in the pages of this volume. Therefore, in closing, we wholeheartedly encourage future research on the full range of psychological mechanisms in animal communication and a greater diversity of species that communicate using signals from a broader range of signaling modalities.

Acknowledgments We thank the series editors, Vincent Janik and Peter McGregor, for their helpful feedback on this chapter. During the preparation of this chapter and while editing the book, MAB was funded by NIH R01 DC009582 and NSF grant IOS-1452831, and CTM was funded by NIH R01 DC012087, R21 MH104756, and NSF IDBR 1254309.

References

Andrews K (2015) The animal mind: an introduction to the philosophy of animal cognition. Routledge, New York

Bee MA (2006) Individual recognition in animal species. In: Naguib M (ed) Encyclopedia of language and linguistics, 2nd edn. Elsevier, New York, pp. 617–626

Beletsky L (1996) The red-winged blackbird: the biology of a strongly polygynous songbird. Academic, San Diego

Beletsky LD, Orians GH (1996) Red-winged blackbirds: decision-making and reproductive success. University of Chicago Press, Chicago

Bizley JK, Cohen YE (2013) The what, where and how of auditory-object perception. Nat Rev Neurosci 14(10):693–707

Bradbury JW, Vehrencamp SL (2011) Principles of animal communication, 2nd edn. Sinauer Associates, Sunderland

Bregman AS (1990) Auditory scene analysis: the perceptual organization of sound. MIT Press, Cambridge

Cheney DL, Seyfarth RM (2010) Primate communication and human language: continuities and discontinuities. In: Kappeler PM, Silk JB (eds) Mind the gap: tracing the origins of human universals. Harvard University Press, Cambridge, pp. 283–298

Dent ML, Bee MA (in press) The principles of auditory object formation in nonhuman animals. In: Slabbekoorn H, Dooling RJ, Popper AN, Fay RR (eds) Human-generated sound and animals. Springer, New York

Doupe AJ, Kuhl PK (1999) Birdsong and human speech: common themes and mechanisms. Annu Rev Neurosci 22:567–631

Fitch WT (2010) The evolution of language. Cambridge University Press, Cambridge

Gerhardt HC, Roberts JD, Bee MA, Schwartz JJ (2000) Call matching in the quacking frog (*Crinia georgiana*). Behav Ecol Sociobiol 48(3):243–251

Godard R (1991) Long-term memory of individual neighbours in a migratory songbird. Nature 350:228–229

Griffiths TD, Warren JD (2004) What is an auditory object? Nat Rev Neurosci 5(11):887–892

Guilford T, Dawkins MS (1991) Receiver psychology and the evolution of animal signals. Anim Behav 42:1–14

Guilford T, Dawkins MS (1993) Receiver psychology and the design of animal signals. Trends Neurosci 16(11):430–436

Hebets EA, Papaj DR (2005) Complex signal function: developing a framework of testable hypotheses. Behav Ecol Sociobiol 57(3):197–214

Hebets EA, Barron AB, Balakrishnan CN, Hauber ME, Mason PH, KL H (2016) A systems approach to animal communication. Proc R Soc Lond B Biol Sci 283(1826):20152889

Janik VM, Slater PJB (1997) Vocal learning in mammals. Adv Study Behav 26:59–99

Kitchen DM (2004) Alpha male black howler monkey responses to loud calls: effect of numeric odds, male companion behaviour and reproductive investment. Anim Behav 67:125–139

Kubovy M, Van Valkenburg D (2001) Auditory and visual objects. Cognition 80(1–2):97–126

Maynard Smith J, Harper D (2003) Animal signals. Oxford University Press, Oxford

Mervis CB, Rosch E (1981) Categorization of natural objects. Annu Rev Psychol 32(1):89–115

Miller CT, Bee MA (2012) Receiver psychology turns 20: is it time for a broader approach? Anim Behav 83(2):331–343

Miller CT, Cohen YE (2010) Vocalizations as auditory objects: behavior and neurophysiology. In: Platt M, Ghazanfar AA (eds) Primate neuroethology. Oxford University Press, Oxford, pp. 236–254

Newen A, Bartels A (2007) Animal minds and the possession of concepts. Philos Psychol 20 (3):283–308

Orians GH, Christman GM (1968) A comparative study of the behaviour of red-winged, tricolored, and yellow-headed blackbirds, vol 84. University of California Publications in Zoology, San Diego

Pearson JM, Watson KK, Platt ML (2014) Decision making: the neuroethological turn. Neuron 82 (5):950–965

Rowe C (2013) Receiver psychology: a receiver's perspective. Anim Behav 85(3):517–523

Ryan MJ, Akre KL, Kirkpatrick M (2009) Cognitive mate choice. In: Dukas R, Ratcliffe JM (eds) Cognitive ecology II. University of Chicago Press, Chicago, pp. 137–155

Scarantino A, Clay Z (2015) Contextually variable signals can be functionally referential. Anim Behav 100:e1–e8

Scholl BJ (2001) Objects and attention: the state of the art. Cognition 80(1–2):1–46

Seyfarth RM, Cheney DL (2009) Seeing who we hear and hearing who we see. Proc Natl Acad Sci USA 106(3):669–670

Seyfarth RM, Cheney DL (2015a) The evolution of concepts about agents: or, what do animals recognize when they recognize an individual? In: Margolis E, Laurence S (eds) The conceptual mind: new directions in the study of concepts. MIT Press, Cambridge, pp. 57–76

Seyfarth RM, Cheney DL (2015b) Social cognition. Anim Behav 103:191–202

Shettleworth SJ (2010) Cognition, evolution, and behavior, 2nd edn. Oxford University Press, Oxford

Soha JA, Peters S (2015) Vocal learning in songbirds and humans: a retrospective in honor of Peter Marler. Ethology 121(10):933–945

Spelke ES, Breinlinger K, Jacobson K, Phillips A (1993) Gestalt relations and object perception: a developmental study. Perception 22(12):1483–1501

Stegmann UE (ed) (2013) Animal communication theory: information and influence. Cambridge University Press, Cambridge

Van Valkenburg D, Kubovy M (2004) From Gibson's fire to Gestalts: a bridge-building theory of perceptual objecthood. In: Neuhoff JG (ed) Ecological Psychoacoustics. Elsevier, Boston, pp. 114–147

Wheeler BC, Fischer J (2012) Functionally referential signals: a promising paradigm whose time has passed. Evol Anthropol 21(5):195–205

Wheeler BC, Fischer J (2015) The blurred boundaries of functional reference: a response to Scarantino and Clay. Anim Behav 100:e9–e13

Wilson M, Hauser MD, Wrangham RW (2001) Does participation in intergroup conflict depend on numerical assessment, range location, or rank for wild chimpanzees? Anim Behav 61(6):1203–1216

Yasukawa K, Searcy WA (1995) Red-winged blackbird (*Agelaius phoeniceus*). Cornell Lab of Ornithology. http://bna.birds.cornell.edu/bna/species/184. Accessed 1 May 2016

Chapter 2

Avian Auditory Processing at Four Different Scales: Variation Among Species, Seasons, Sexes, and Individuals

Kenneth S. Henry, Megan D. Gall, Alejandro Vélez, and Jeffrey R. Lucas

Abstract Previous research on songbirds has typically focused on variation in production of vocal communication signals. These studies have addressed the mechanisms and functional significance of variation in vocal production across species and, within species, across seasons and among individuals (e.g., males of varying resource-holding capacity). However, far less is known about parallel variation in sensory processing, particularly in non-model species. A relationship between vocal signals and auditory processing is expected based on the sender–receiver matching hypothesis. Here, we review our recent comparative studies of auditory processing in songbirds conducted using auditory evoked potentials (AEPs) in a variety of field-caught songbird species. AEPs are voltage waveforms recorded from the scalp surface that originate from synchronous neural activity and provide insight into the sensitivity, frequency resolution, and temporal resolution of the subcortical auditory system. These studies uncovered variation in auditory processing at a number of different scales that was generally consistent with the sender–receiver matching hypothesis. For example, differences in auditory processing were uncovered among species and across seasons that may enhance perception of communication signals in species-specific habitats and during periods of mate selection, respectively. Sex differences were also revealed, often in season-specific patterns, and surprising individual differences were observed in auditory

---

K.S. Henry
Department of Otolaryngology, University of Rochester, Rochester, NY, USA
e-mail: kenneth_henry@urmc.rochester.edu

M.D. Gall
Department of Biology, Vassar College, Poughkeepsie, NY, USA
e-mail: megall@vassar.edu

A. Vélez
Department of Biology, Washington University, St. Louis, MO, USA
e-mail: avelezmelendez@wustl.edu

J.R. Lucas (✉)
Department of Biological Sciences, Purdue University, W. Lafayette, IN, USA
e-mail: jlucas@purdue.edu

M.A. Bee, C.T. Miller (eds.), *Psychological Mechanisms in Animal Communication*, Animal Signals and Communication 5,
DOI 10.1007/978-3-319-48690-1_2

processing of mate attraction signals but not of calls used in interspecific contexts. While much remains to be learned, these studies highlight a previously unrecognized degree of parallel variation in songbirds, existing at diverse hierarchical scales, between production of vocal communication signals and subcortical auditory processing.

2.1 Introduction

The oscine passerines, or songbirds, are an important system for studying animal communication because, as a group, they rely heavily on vocalizations for survival and reproduction. A considerable body of research has focused on the mechanisms and adaptive significance of variation in vocal production within this group. At one level, considerable variation exists in the acoustic structure and complexity of species-specific vocalizations across approximately 4000 extant species. Vocal variation across species reflects differences in underlying brain circuitry and vocal tract anatomy (Zeng et al. 2007) and may ultimately enhance the fidelity of information transfer in species-specific habitats (Morton 1975). At a second level of variation, differences in signal production exist between individuals of the same species. For example, production of songs (mate attraction signals) in species inhabiting temperate latitudes typically occurs in males but not females and may vary considerably among males in relation to their resource-holding potential (e.g., Christie et al. 2004). A third level of variation in vocal production occurs within individuals over time, that is, during development and across seasons. Male songbirds of many species show dramatic seasonal changes in song production throughout the year that appear functionally adaptive and are linked to hormonally mediated regulation of underlying song control nuclei (Brenowitz 2004).

In contrast to our relatively extensive knowledge of vocal production in songbirds, far less is known about variation in auditory processing, particularly in non-model species. Indeed, the auditory capabilities of receivers in animal communication systems are often implicitly assumed to match our own perceptual abilities or perhaps spectral properties of vocal signals reflected in a spectrogram. In this chapter, we summarize our recent studies of the mechanisms and adaptive significance of auditory variation in songbirds. These studies have extended our knowledge of animal communication systems by examining auditory processing in songbirds primarily at the three aforementioned levels of variation: (1) auditory variation among species, (2) variation among individuals of the same species (e.g., between sexes), and (3) seasonal variation in auditory processing operating at the level of the individual.

The broad scale of these investigations was made possible through use of an efficient physiological method for evaluating subcortical auditory processing known as *auditory evoked potentials (AEPs)*. AEPs are average voltage waveforms recorded from the scalp in response to acoustic stimulation that arise from synchronized activity within neural populations located along the auditory pathway (Hall 1992). Only synchronous discharges from sufficiently large populations of neurons produce a large enough gross potential to be measurable at the scalp surface. While the gross nature of AEPs provides minimal information about the response properties of single neurons and small populations of neurons, their ability to provide basic insight into the absolute sensitivity, frequency resolution, and temporal resolution of the subcortical auditory system has led to recent increases in the application of this technique.

Processing of acoustic information by the auditory system has traditionally been measured by recording electrical activity from single neurons or small groups of neurons located in nuclei of the ascending auditory pathway in anesthetized, nonhuman animals (e.g., Konishi 1970; Chap. 3). One of the most fundamental aspects of auditory processing revealed by these studies is the tonotopic representation of sound, that is, different frequency components of the acoustic spectrum are represented by activity in different frequency-tuned neural subpopulations or channels. Spectral decomposition of sounds into their frequency components originates through auditory filtering in the inner ear and is maintained to varying degrees in different auditory nuclei along the entire length of the pathway from the brainstem nuclei to the midbrain, thalamus, and forebrain (e.g., Calford et al. 1983). A second, fundamental aspect of processing is neural synchrony to the temporal structure of sound. The instantaneous firing rate of auditory neurons varies with the amplitude envelope of acoustics signals (envelope following; Joris et al. 2004) and, at more peripheral levels of processing, the fine structure of the pressure waveform as well (typically for frequencies less than 3–6 kHz; frequency following; Johnson 1980). Envelope following occurs along the entire pathway up to and including the auditory forebrain, with the maximum frequency of encoded envelope fluctuations decreasing centrally (Joris et al. 2004). Synchrony of neural responses to envelope fluctuations, fine structure, and particularly sudden onsets of sound is the key aspect of auditory processing that allows assessment of auditory function through AEPs.

2.2 Assessment of Hearing Function with AEPs

AEPs are obtained by averaging scalp-recorded voltage waveforms across a large number of stimulus repetitions (typically several hundred to several thousand). Recordings are conducted in anesthetized birds to minimize muscle artifacts and with acoustic stimuli presented either from an electromagnetically shielded, free-field loudspeaker or insert earphone in a sound-attenuating booth. Subcutaneous electrodes are positioned at the vertex of the skull (non-inverting), posterior to the ipsilateral ear canal (inverting), and the nape of the neck (common ground) for recording AEPs. Voltage signals are differentially amplified with a gain of 100–200 K and band-pass filtered from between 0.1 and 5–10 kHz (depending on the measurement of interest). Sufficient electromagnetic shielding is critically important to prevent contamination of AEPs by stimulus artifact and can be verified with recordings made with the electrodes in saline or a potato. Detailed AEP recording procedures are described in a number of publications from our laboratory (e.g., Henry and Lucas 2008; Gall et al. 2013; Lucas et al. 2015; Vélez et al. 2015a).

AEP waveforms recorded in response to sound show a fast, bipolar deflection in response to the onset of acoustic stimulation (Fig. 2.1a) known as the *auditory brainstem response (ABR)*. ABRs in response to click stimuli and tone bursts with fast onset ramps contain multiple peaks attributable to different neural generators of the auditory pathway (reviewed in Hall 1992). The shortest latency peak is generated by the auditory nerve, while peaks with greater latencies arise from nuclei located further along the auditory pathway. The amplitude of ABR peaks varies with the frequency spectrum and sound level of the stimulus in relation to the number of underlying, responsive neurons and their synchrony. The latency of ABR peaks varies with sound level, with louder stimuli evoking shorter latency responses, and to a lesser extent with frequency (higher frequency stimuli tend to evoke ABRs of shorter latency). For responses to sounds with durations longer than a click, the ABR peaks are followed by a sustained response that reflects neural synchrony to the temporal fine structure and envelope of the stimulus waveform. The component of the sustained response that follows the stimulus envelope (Fig. 2.1b) is known as the *envelope following response (EFR)*, while the component locked to the low-frequency fine structure (Fig. 2.1a, b) is called the *frequency following response (FFR)*.

The audiogram, which plots absolute threshold (i.e., minimum detectable sound pressure level in quiet) as a function of frequency, delineates the frequency range of sensitive hearing and, therefore, serves as a starting point for understanding the auditory capabilities of any species. Audiograms can be estimated from ABRs recorded in response to tone burst stimuli of varying frequency and sound pressure level. For each tone frequency, the threshold is estimated as the minimum sound pressure level that evokes a reliable ABR. Thresholds are traditionally assessed by visual inspection of ABR waveforms but can also be estimated using statistical techniques such as linear regression or signal detection theory (see Gall et al. 2011). Compared to absolute thresholds of single neurons and of behaving animals, thresholds based on ABRs typically show a good correlation across frequencies and are elevated by 20–30 dB (Brittan-Powell et al. 2002).

As previously discussed, neurons of the auditory system are arranged into frequency-tuned, tonotopic channels that span the auditory pathway. The frequency bandwidths of these auditory filters are an important parameter of the system because these bandwidths determine the extent to which two sound components of similar frequency stimulate different neural populations, that is, are resolved by the auditory system. While the gross nature of AEPs would appear to preclude the possibility of measuring frequency resolution, estimates of auditory filter bandwidths can be obtained by combining the tone-evoked ABR methodology with the notched noise paradigm of human psychophysics (Patterson 1976). Notched noise is white noise with a narrow frequency band of energy filtered out around a center frequency. In traditional notched noise experiments (reviewed in Moore 1993),

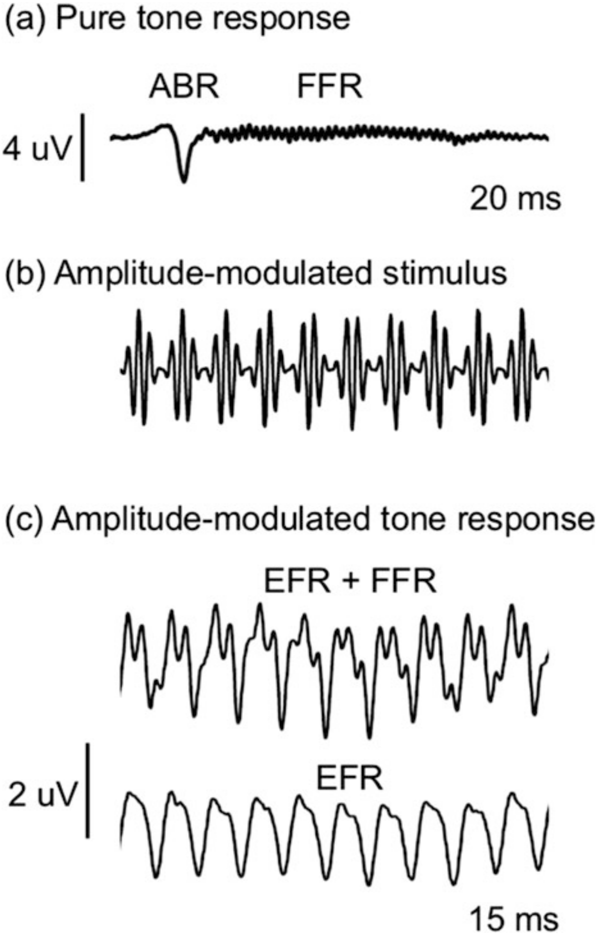


Fig. 2.1 Examples of auditory evoked potentials (AEPs). (**a**) An AEP recorded in response to a 3-kHz pure tone with 3-ms onset and offset ramps. The response consists of an auditory brainstem response (ABR) at stimulus onset followed by a sustained frequency following response (FFR) associated with auditory temporal coding of the stimulus fine structure. (**b**, **c**) Responses to an amplitude-modulated tone stimulus. (**b**) Amplitude-modulated tone stimulus waveform with a carrier frequency of 2.75 kHz and an amplitude modulation frequency of 710 Hz and (**c**) AEP responses. The stimulus waveform plots pressure (arbitrary scale) as a function of time. The portion of the AEP responses to amplitude modulation shown in (**c**) begins 15 ms after stimulus onset and hence excludes the ABR. The *top* trace in (**c**) contains both the FFR (auditory temporal coding of the carrier frequency and amplitude modulation sidebands) and the envelope following response (EFR; auditory temporal coding of the stimulus modulation frequency). The lower trace in (**c**) has been low-pass filtered at 1 kHz to remove the FFR and hence contains only the EFR component. AEPs in (**a**) and (**c**) were recorded from a dark-eyed junco (*Junco hyemalis*) and house sparrow (*Passer domesticus*), respectively. The total durations of the waveforms are given in the *lower right corners* of panels (**a**) and (**c**)

behavioral thresholds for detection of the tone are measured in notched noise across a range of notch bandwidths. The center frequency of the notch is typically held constant at the frequency of the tone signal, but can be varied to investigate filter asymmetry. Auditory filter shapes are derived from these data based on three main assumptions: (1) the tone is processed in a single auditory filter channel, (2) only noise energy that falls into that auditory filter channel contributes to masking the tone, and (3) detection of the tone requires a fixed signal-to-noise ratio at the output of the filter. When the notch bandwidth is sufficiently broad, little noise energy enters the auditory filter and the threshold for signal detection is low. When the notch bandwidth is narrower than the auditory filter, however, substantial noise energy spills into the filter and contributes to masking of the signal, resulting in a relatively higher threshold. In our adaptation of this method, auditory filter shapes are derived using ABR threshold data collected across a range of notch bandwidths fit to Patterson's rounded exponential auditory filter model (Fig. 2.2). While the precise relationship between ABR-based tuning bandwidths and the frequency tuning of single neurons has not been quantified in any one species, the ABR method produces estimates of auditory filter bandwidth that are reasonable based on comparison to single-neuron data from other species.

Within each tonotopic channel of the auditory system, the amplitude envelope of sound after auditory filtering is encoded through oscillations in the instantaneous firing rate of auditory neurons (Joris et al. 2004). The ability of auditory neurons to follow fast envelope fluctuations is a second, important parameter of the system because it determines the extent to which sounds occurring in rapid succession are temporally resolved versus fused together during auditory processing. In one method for assessing temporal resolution with AEPs, EFRs are recorded in response to amplitude-modulated sounds to generate a *modulation transfer function (MTF)* plotting EFR amplitude (i.e., the amplitude of the spectral component of the response at the modulation frequency of the stimulus) as a function of modulation frequency (Kuwada et al. 1986; Dolphin and Mountain 1992; Schrode and Bee 2015). EFR-based MTFs in birds typically show a peak at modulation frequencies from 300 to 700 Hz followed by steady declines in EFR amplitude at higher modulation frequencies attributable to reduced fidelity of neural envelope coding (e.g., Henry and Lucas 2008). Differences in EFR amplitude at high modulation frequencies can reflect differences in the temporal resolution of the auditory system.

In a second method for quantifying temporal resolution, ABRs are recorded in response to paired click stimuli presented with a brief time interval inserted between clicks (Supin and Popov 1995; Parham et al. 1996; Ohashi et al. 2005; Schrode and Bee 2015). ABR recovery functions generated from paired click responses plot the amplitude of the response to the second click, usually expressed as a percentage of amplitude of the response to a single click, as a function of the duration of the time interval between clicks. Because the amplitude of the ABR to the second click can be difficult to measure for short inter-click intervals, the ABR to the second click is usually derived by waveform subtraction of a single-click response from the response to both clicks (Fig. 2.3). ABR recovery functions in a variety of taxa, including birds, show a steady reduction in the amplitude of the

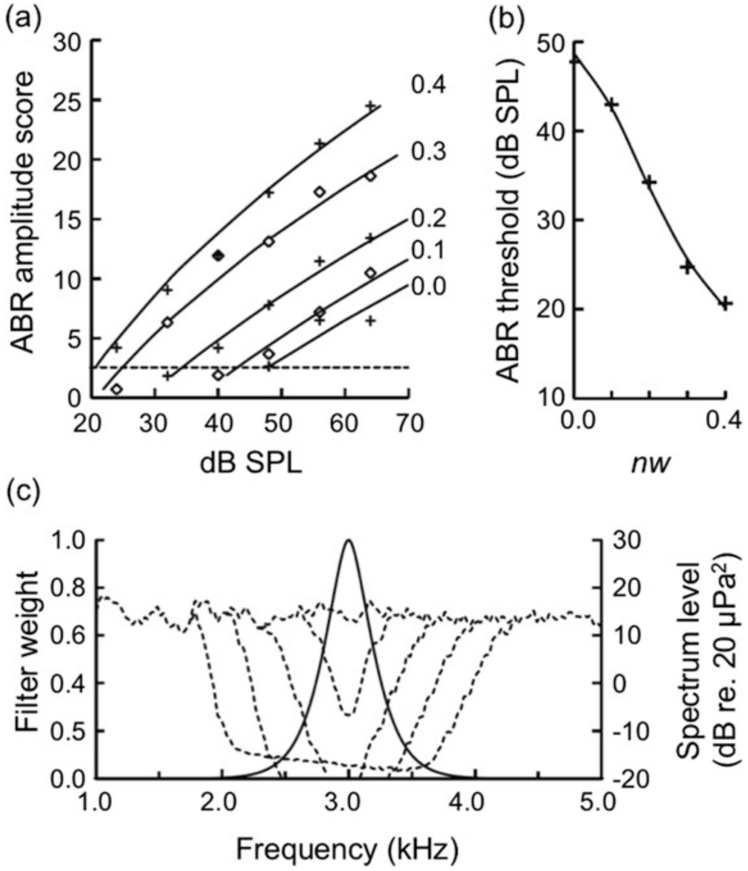


Fig. 2.2 Estimation of auditory filter bandwidth from tone-evoked ABRs in notched masking noise (from Henry and Lucas 2010a [Fig. 1]). The frequency of tone stimuli was 3 kHz. (**a**) ABR amplitude (normalized) plotted as a function of stimulus level at five different notch bandwidths [*nw* (*right*); expressed as the bandwidth in Hz divided by twice the center frequency of the notch (2 × 3 kHz)]. ABRs were recorded from a tufted titmouse (*Baeolophus bicolor*). *Solid lines* show the fit of a generalized linear model used for estimation of ABR thresholds. The ABR threshold for each notch bandwidth is the sound pressure level (dB SPL) at which the ABR amplitude score exceeds a statistical criterion (*dashed line*). (**b**) ABR thresholds (+ symbols) from panel (**a**) plotted as a function of notch bandwidth. The *solid line* represents the fit of the *roex* (p, r) auditory filter model ($p = 23.4$, $r = 0.000030$, $K' = 41.7$). (**c**) The shape of the auditory filter derived from the threshold function in (**b**) (*solid line*; left y-axis). Also drawn are the long-term average power spectra (bin width = 25 Hz) of the notched noise maskers (*dashed lines*; right y-axis)

ABR to the second click with decreasing inter-click interval associated with a reduction in the neural representation of the second click. Differences in ABR recovery observed at relatively short inter-click intervals can reflect differences in the temporal resolution of the subcortical auditory system.

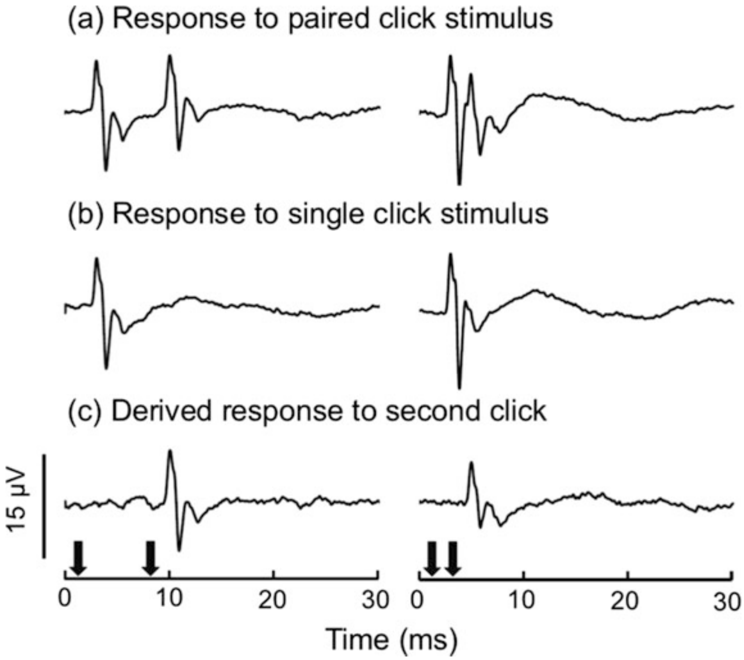


Fig. 2.3 Estimation of auditory temporal resolution using ABRs evoked by paired-click stimuli (from Henry et al. 2011 [Fig. 2]). Shown here are ABRs in response (**a**) to paired-click stimuli and (**b**) to single-click stimuli. (**c**) Derived ABRs to the second click of paired stimuli, which were generated by point-to-point subtraction of the response to the single-click stimulus from the response to the paired-click stimulus. *Thick arrows* indicate the timing of clicks, which were separated by time intervals of 7 ms (*left*) and 2 ms (*right*)

While the ABR is relatively straightforward to interpret due to its position in time near sound onset, EFRs and FFRs can present greater challenges. For example, sustained responses to the same stimulus component from different neural generators may combine to varying degrees either constructively or destructively depending on their relative amplitudes, the difference in response latency between generators, and the frequency of the stimulus component being followed. The problem is mitigated to some extent by the tendency for one neural generator to dominate a particular class of AEPs (e.g., EFRs to amplitude-modulated sounds may be dominated by an auditory nerve component at modulation frequencies greater than a few hundred Hz; Henry and Lucas 2008), but exists nonetheless. The depth of anesthesia can also affect auditory processing, either directly, as is the case for relatively central auditory nuclei, or through an intermediate effect on body temperature. Care should be taken to maintain stable body temperature and consistent depth of anesthesia both within and across AEP recording sessions. In our studies, we routinely used click-evoked ABRs to assess the stability of auditory function during recording sessions.

2.3 Coevolution Between Signalers and Receivers

Vocal communication signals are very diverse and often mediate evolutionarily important behaviors such as species recognition, mate attraction, territory defense, and group cohesion (Kroodsma and Miller 1996; Marler and Slabbekoorn 2004). Therefore, natural and sexual selection are expected to shape vocal signals in ways that optimize the transfer of information from signalers to receivers (Bradbury and Vehrencamp 2011). Because communication errors can have costly fitness consequences, signal-processing mechanisms in receivers are expected to match the physical properties of communication signals (Endler 1992). This expectation has been referred to as the matched filter hypothesis (Capranica and Moffat 1983) or the sender–receiver matching hypothesis (Dooling et al. 2000; Woolley et al. 2009; Gall et al. 2012a). In this section, we review behavioral and physiological studies looking at the match between signal properties and auditory processing in birds.

The correlation between vocal spectral content and the frequency range of auditory sensitivity (the audiogram) is arguably the most studied dimension of the sender–receiver matching hypothesis in birds. Konishi (1970) obtained auditory thresholds from single units in the cochlear nucleus of ten species of songbirds spanning six families: Emberizidae (five species), Icteridae (one species), Turdidae (one species), Sturnidae (one species), Passeridae (one species), and Fringillidae (one species). Species differences in high-frequency hearing sensitivity appear to be correlated with differences in the range of frequencies present in the conspecific vocal repertoire. These results offer support for the sender–receiver matching hypothesis. However, the author noted that covariation between vocal frequency content and the range of high-frequency hearing can also be due to differences in body size. The size of sound-producing structures can impose limits on the frequencies that songbirds can produce (Greenewalt 1968). Similarly, the size of the tympanic membrane and middle ear bone (columella in birds) can impose limits on the frequency range of hearing. Interestingly, however, canaries (Fringillidae) and house sparrows (Passeridae) have lower high-frequency sensitivity than white-crowned sparrows and white-throated sparrows (Emberizidae), even though they are very similar in size. Given the similarities between the spectral content of the vocalizations among these four species, differences in auditory sensitivity do not reflect a match between signalers and receivers. Although signal properties and body size fail to explain differences in auditory sensitivity in these species, it is possible that auditory processing differences are due to taxonomic differences. In fact, emberizids were overall more sensitive to high-frequency sounds than all other species, suggesting phylogenetic conservation in auditory processing.

Behavioral studies using psychophysical methods have also been used to examine the relationship between hearing and vocal performance. Dooling (1992) reviews these studies, but some patterns are worth noting here. In general, high-frequency hearing sensitivity correlates with the high-frequency components in species-specific vocalizations (Dooling et al. 1978; Dooling 1982). Interestingly, Dooling (1992) found that emberizids are more sensitive to high-frequency sounds than all other families covered in his review. These results are in line with those of Konishi (1970) and suggest that phylogenetic conservation can constrain the evolution of auditory processing mechanisms.

More recently, studies using AEPs to measure peripheral auditory processing have also shown correlations between frequency sensitivity and spectral vocal content. Henry and Lucas (2008) showed that tufted titmice (*Baeolophus bicolor*), house sparrows (*Passer domesticus*), and white-breasted nuthatches (*Sitta carolinensis*) have similar hearing thresholds for sounds with frequencies below 4 kHz (Fig. 2.4). However, ABR thresholds to tones of 6.4 kHz were 12–14 dB lower (more sensitive) in tufted titmice than the other three species. These results were interpreted as an adaptation for processing high-frequency alarm calls present in the vocal repertoire of tufted titmice. Following up on these results, Lucas et al. (2015) measured the minimum, peak, and maximum frequencies in calls and songs of these species and of the white-crowned sparrow (*Zonotrichia leucophrys*). Overall, white-crowned sparrows have higher minimum, peak, and maximum frequencies than all other species. Accordingly, Lucas et al. (2015) also showed that white-crowned sparrows are less sensitive below 4 kHz and more sensitive above 4 kHz than the other species (Fig. 2.4). These results are in line with predictions of the sender–receiver matching hypothesis.

The sender–receiver matching hypothesis is not restricted to the frequency range of vocal signals and frequency sensitivity. Vocalizations also vary in temporal structure and the temporal properties of vocal signals can play important roles in avian communication (Beckers and TenCate 2001). Gall et al. (2012a) investigated

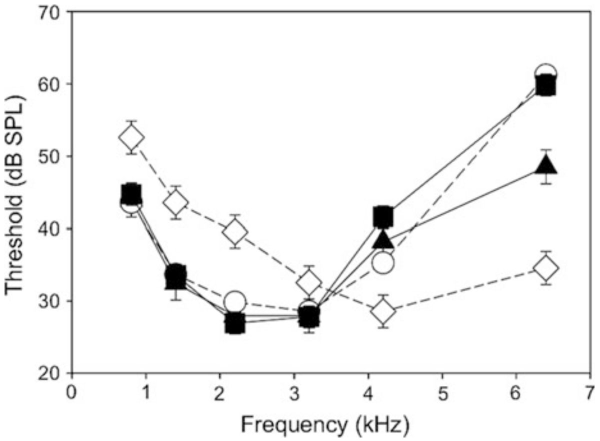


Fig. 2.4 ABR-based audiograms of four songbird species (from Lucas et al. 2015 [Fig. 5]). Audiograms plot tone-evoked ABR thresholds (LS means ± s.e.m.) as a function of stimulus frequency for titmice (*Baeolophus bicolor*; *triangles*), nuthatches (*Sitta carolinensis*; *squares*), house sparrows (*Passer domesticus*; *circles*), and white-crowned sparrows (*Zonotrichia leucophrys*; *diamonds*). Thresholds were estimated using a visual detection method. Woodland species, *closed symbols* and *continuous lines*; open-habitat species, *open symbols* and *dashed lines*

whether auditory processing of the rise time (or onset time) and of the sustained portion of sounds varies according to spectro-temporal features of vocalizations in five species of songbirds. Song elements of brown-headed cowbirds (*Molothrus ater*) have the most rapid rise times, followed by songs of dark-eyed juncos (*Junco hyemalis*), and, with the slowest rise times, song elements of white-crowned sparrows, American tree sparrows (*Spizella arborea*), and house finches (*Carpodacus mexicanus*). Based on these song properties, the authors predicted strong ABRs to the onset of sounds in brown-headed cowbirds, followed by dark-eyed juncos, and with the weakest and similar ABRs, all other species in the study. Accordingly, ABRs to the onset of sounds were stronger in dark-eyed juncos than in white-crowned sparrows, American tree sparrows, and house finches. Contrary to predictions, however, brown-headed cowbirds had the weakest onset ABRs. Regarding the tonality of songs, white-crowned sparrows, American tree sparrows, and house finches also had more tonal song elements with the slowest rates of frequency modulation, followed by dark-eyed juncos, and with the fastest rates of frequency modulation and less tonal song elements, brown-headed cowbirds. Therefore, the authors predicted the strongest FFRs to the sustained portion of sounds would be found in white-crowned sparrows, American tree sparrows, and house finches, because they have more tonal song elements, followed by dark-eyed juncos, with the weakest FFRs found in brown-headed cowbirds, because they have less tonal song elements. As predicted, FFRs to the sustained portion of the sound were weak in brown-headed cowbirds. However, FFRs were strongest in dark-eyed juncos.

With some exceptions, these results largely support the sender–receiver matching hypothesis. Despite having relatively high-frequency vocalizations with rapid onsets, brown-headed cowbirds have weak ABRs to the onset of sounds, particularly at higher frequencies (Gall et al. 2011, 2012b). One possible explanation for the mismatch between signal properties and auditory processing in brown-headed cowbirds relates to their breeding strategy (Gall et al. 2011, 2012b). Brown-headed cowbirds are brood parasites, and processing of heterospecific vocalizations may therefore be very important to locate appropriate hosts. In fact, as discussed below, Gall and Lucas (2010) report drastic sex differences in auditory filter bandwidth in brown-headed cowbirds that align with processing of heterospecific vocalizations in females. Another possibility for the mismatch between signalers and receivers in brown-headed cowbirds is phylogenetic conservation. Brown-headed cowbirds and red-winged blackbirds (*Agelaius phoeniceus*) belong to the family Icteridae and share great similarities in auditory processing (Gall et al. 2011). Phylogenetic conservation could also explain why dark-eyed juncos and American tree sparrows, both members of the Emberizidae family, had similar ABRs and FFRs, despite differences in their vocalizations. However, white-crowned sparrows, another species in the Emberizidae family, had ABRs and FFRs more similar to brown-headed cowbirds and house finches.

Avian vocalizations often have more than one frequency component. For instance, the vocal repertoire of zebra finches (*Taeniopygia guttata*) includes harmonic complexes with over 15 frequency components. Differences in the frequency separation and the relative amplitude of each component lead to differences in pitch and timbre. Interestingly, the interaction between frequency components generates a gross temporal structure (Bradbury and Vehrencamp 2011) that may be processed in the temporal domain (Hartmann 1998). Through psychophysical experiments, Lohr and Dooling (1998) showed that zebra finches and budgerigars (*Melopsittacus undulatus*) are better than humans at detecting changes in one frequency component of complex harmonic stimuli. Furthermore, they showed that zebra finches perform better than budgerigars, which have predominantly frequency-modulated tonal vocalizations with less complex temporal structure. These results offer support for the sender–receiver matching hypothesis in that birds that produce structurally complex vocalizations are better at processing the temporal structure of complex sounds than birds that produce simple, tonal vocalizations.

AEPs have also been used to investigate peripheral auditory processing of temporal properties of sounds in the context of the sender–receiver matching hypothesis. Henry and Lucas (2008) measured the rates of envelope fluctuation in vocalizations of tufted titmice, house sparrows, and white-breasted nuthatches. Envelope fluctuation rates of vocalizations were higher in house sparrows and tufted titmice, with rates as high as 1450 Hz. Accordingly, the EFR to amplitude-modulated tones was stronger in house sparrows and tufted titmice than in nuthatches at modulation rates above 1 kHz. These results suggest that the auditory system of species with temporally structured vocalizations has higher temporal resolution. It is important to note that the song of white-breasted nuthatches is a harmonic complex with strong amplitude fluctuations at rates around 700 Hz. Vélez et al. (2015b) showed that EFRs to tones with envelope fluctuations between 200 and 900 Hz were stronger in white-breasted nuthatches during the spring than in two species with tonal songs: Carolina chickadees (*Poecile carolinensis*) and tufted titmice.

To conclude, there is ample evidence to support a general match between vocal properties and auditory processing in songbirds. This match, however, can be constrained by gross morphological and physiological differences between species, ultimately due to more distant phylogenetic relationships. For this reason, comparisons across a few distantly related species may produce biased results, and the observed correlations between properties of the vocalizations and the auditory system may not be solely due to coevolution between signalers and receivers. Therefore, we propose that comparative studies of closely related species that take into account different factors potentially shaping the evolution of vocal signals are an important next step to better understand whether and how auditory processing mechanisms and signal properties coevolve. Similarly, studies of variation in auditory processing across populations of the same species that differ in song properties (e.g., Slabbekoorn and Smith 2002; Derryberry 2009) might shed additional light on the correlation between signal design and receiver sensory physiology.

2.4 Habitat Effects on Song and Hearing

The acoustic adaptation hypothesis proposes that habitat structure constrains the evolution of acoustic signals (Morton 1975). In forests, reverberations and excess attenuation of high-frequency sounds favor tonal signals with frequencies below 3–5 kHz. In open habitats, the lack of reverberations and slow amplitude fluctuations imposed by wind select for high-frequency acoustic signals rapidly modulated in amplitude, frequency, or both (Morton 1975; Marten and Marler 1977; Wiley and Richards 1978; Richards and Wiley 1980; Wiley 1991). Birds have long been used as a model system to test the acoustic adaptation hypothesis. Results from these studies often agree with predictions from the acoustic adaptation hypothesis and show that species that live in forests have more tonal songs with lower frequencies than species that live in open habitats (reviewed in Boncoraglio and Saino 2007). Based on the sender–receiver matching hypothesis, habitat-dependent differences in song properties could lead to habitat-related differences in auditory processing of tonal versus temporally structured sounds and low-frequency versus high-frequency sounds.

We can use auditory filter properties as an index of the trade-off between processing tonal versus temporally structured sounds. Narrow filters have greater spectral resolution, whereas broad filters have greater temporal resolution (Henry et al. 2011). Because selection favors vocal signals that are tonal in forests and amplitude-modulated in open habitats, Henry and Lucas (2010b) predicted narrower auditory filters in forest species than in open-habitat species. As predicted, they found that auditory filters were generally narrower in forest species than open-habitat species (Fig. 2.5). Auditory filter bandwidth was significantly lower in white-breasted nuthatches, a forest species, than in house sparrows and white-crowned sparrows, both open-habitat species. Tufted titmice, another forest species, had significantly narrower auditory filter bandwidths compared to house sparrows, but not compared to white-crowned sparrows.

Interestingly, auditory filter properties of dark-eyed juncos, another forest species, differed drastically from those of all other species. While auditory filter bandwidth increased with increasing frequency in all other species, auditory filter bandwidth remained constant across test frequencies from 2 to 4 kHz in dark-eyed juncos. This result is striking for two main reasons. First, auditory filters with constant bandwidth across frequencies are uncommon and tend to occur in auditory specialists like barn owls (*Tyto alba*; Köppl et al. 1993) and echolocating species (Suga et al. 1976; Popov et al. 2006). Second, the pattern differs from those of the other forest species and from that of white-crowned sparrows, the most closely related species to dark-eyed juncos in the study. These results suggest that factors other than habitat and phylogenetic relatedness between species may have shaped auditory filter properties in dark-eyed juncos. Henry and Lucas (2010b) also found that the response of the auditory filters was more efficient (measured as the signal-to-noise response threshold) in open-habitat species than in forest species. This

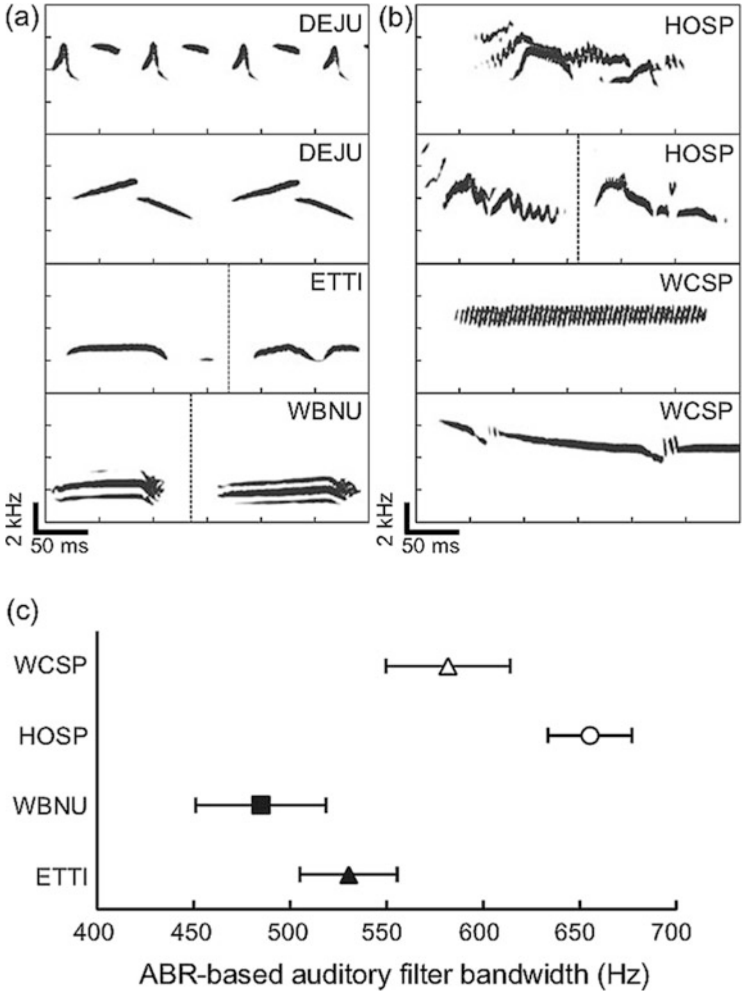


Fig. 2.5 Acoustic structure of long-range vocalizations (songs) and ABR-based estimates of auditory filter bandwidth (frequency resolution) in four songbird species (from Henry and Lucas 2010b [Figs. 4 and 5]). Spectrograms of song notes from (**a**) three woodland species (dark-eyed junco, DEJU; tufted titmouse, ETTI; white-breasted nuthatch, WBNU) and (**b**) two open-habitat species (house sparrow, HOSP; white-crowned sparrow, WCSP). Song notes separated by a *dashed line* within the same panel are different examples. (**c**) Average auditory filter bandwidths among four of the study species while controlling for the effect of frequency. Data points represent least squares means ± s.e.m. of the species effect. Auditory filter bandwidths of the dark-eyed junco (not shown) are similar to the open-habitat species from 2 to 3 kHz and lower than the woodland species at 4 kHz

result was interpreted as a possible adaptation to compensate for the inherently lower sensitivity in noise of broad auditory filters.

Lucas et al. (2015) followed up on this study by examining both the properties of the vocalizations and the gross and fine temporal processing of complex tones in two forest species (white-breasted nuthatches and tufted titmice) and two open-habitat species (house sparrows and white-crowned sparrows). Consistent with the acoustic adaptation hypothesis, they showed that, overall, songs of open-habitat species have higher levels of entropy and higher peak and maximum frequencies than forest species. Entropy is a measure of the relative tonality of a signal, with low values representing tonal sounds and high values representing noisy sounds. One result was inconsistent with the acoustic adaptation hypothesis: white-breasted nuthatch songs are composed of a series of harmonics (Ritchison 1983) that generate a strong amplitude modulation[1] in this forest-adapted species. Moreover, the properties of call notes of both woodland species were inconsistent with the prediction of low-frequency, tonal vocalizations in forest habitats. However, calls are commonly produced in winter when the assumption of propagation-induced sound degradation is less valid due to the lack of leaves.

Based on differences in vocalizations and in auditory filter bandwidth (Henry and Lucas 2010b), Lucas et al. (2015) predicted that the auditory system of open-habitat species should respond stronger to the amplitude envelope of complex tones, while the auditory system of forest species should respond stronger to the individual spectral components of complex tones. They measured EFRs and FFRs evoked by complex tones of two and three frequency components with a fundamental frequency of 600 or 1200 Hz. As predicted, EFRs to the 600 Hz amplitude modulation rate of a complex tone were stronger in open-habitat species than in forest species. Interestingly, differences within forest species show that EFRs were stronger in nuthatches than in tufted titmice when the complex tone had an AM rate of 600 Hz. This more closely resembled nuthatch vocalizations. In contrast, FFRs to the different spectral components of the complex tones varied little across habitats. Indeed, FFRs were strongest in white-crowned sparrows. These results cannot be explained by differences in overall auditory sensitivity. Audiograms of these four species show that white-crowned sparrows are actually less sensitive to sounds in the frequency range of the complex tones used in the experiment (Lucas et al. 2015). Strong EFRs and FFRs in white-crowned sparrows may be due to greater acoustic complexity in their songs, as discussed below.

One surprising result was the potential for forest species to process harmonic stacks in two ways. One way is by processing of the envelope fluctuations of complex tones, as described by the envelope following response above. The second is spectral enhancement where the processing of the tonal properties of a given harmonic is enhanced when that harmonic is coupled with the next lowest harmonic in the series. Nuthatches have a simple vocal repertoire that includes a call and a song with strong harmonic content. Thus, enhanced processing of the individual spectral components and the envelope fluctuations of sounds may be necessary to

---

[1] We use the term “amplitude modulation” here to refer to the periodic modulation, sinusoidal or otherwise, of the temporal amplitude envelope of a signal.

decode all of the information in conspecific vocalizations, as suggested by Vélez et al. (2015a).

These studies, together with those on the sender-receiving matching hypothesis, show that comparative studies that take into account different factors that can shape the evolutionary design of vocal signals are necessary to better understand the evolution of auditory processing mechanisms. In a comparative study with nine species of New World sparrows, Vélez et al. (2015a) investigated whether auditory sensitivity to the frequency of sounds depends on song frequency content, song structure, or habitat-dependent constraints on sound propagation. The selection of species included three that predominantly live in forests, three in scrub-like habitats, and three in open habitats. Within each habitat, one species produces songs that are simple and tonal, one produces trilled songs composed of one element repeated throughout the entire song, and one produces complex songs that include tones, trills, and buzzes. Importantly, more closely related species in the study do not occupy similar habitats nor do they have structurally similar songs. Consistent with the acoustic adaptation hypothesis, Vélez et al. (2015a) found that songs of forest species are more tonal and have lower frequencies than songs of species that live in open habitats. Interestingly, species from different habitats had very similar hearing, as evidenced by audiograms obtained with ABRs. High-frequency hearing sensitivity, however, differed between species with different song types. Species that produce complex songs were more sensitive to high-frequency sounds than all other species. Why auditory sensitivity correlates with the structure of the songs is an open question. One possibility is that birds with complex songs, like the white-crowned sparrow, utilize a broader range of frequencies in order to decode all of the note types in the songs. Thus, the amount of information that is encoded in songs may correlate with auditory sensitivity. These results highlight the importance of considering the multiple dimensions of signals, and how the different dimensions interact, when studying the evolution of signal-processing mechanisms.

Studies of within-species variation are fundamental for understanding evolutionary processes. Several studies have shown how habitat-dependent constraints on sound propagation lead to divergence in song properties across conspecific populations inhabiting different environments (reviewed in Slabbekoorn and Smith 2002; Derryberry 2009). More recently, growing conservation efforts set the stage for studies on the effect of urbanization and anthropogenic noise in animal communication. Gall et al. (2012c) showed that urbanization, as well as habitat, affects the active space of brown-headed cowbird song. Furthermore, differences in the acoustic properties of bird vocalizations between urban and rural conspecific populations have been reported (Slabbekoorn and Peet 2003; Ríos-Chelén et al. 2015). Despite the breadth of studies looking at population-level variation in signal properties, differences in auditory processing between conspecific populations have been overlooked. To better understand the evolution of auditory processing mechanisms, future studies should focus on how auditory processing correlates with habitat-dependent and noise-dependent differences in vocal properties within species.

2.5 Seasonal Auditory Plasticity

As discussed earlier, auditory processing, particularly at the periphery, is often assumed to be static. However, an emerging body of work suggests that there is actually considerable plasticity in the auditory periphery of birds (Lucas et al. 2002, 2007; Henry and Lucas 2009; Caras et al. 2010; Gall et al. 2013; Vélez et al. 2015b), as well as fish (Sisneros et al. 2004; Vasconcelos et al. 2011; Coffin et al. 2012) and anurans (Gall and Wilczynski 2015).

Peripheral auditory processing was first shown to be influenced by season, and consequently by the likely reproductive state of birds, in 2002. Lucas et al. (2002) used AEPs (specifically the ABR) to investigate the response to clicks by five species (downy woodpeckers (*Picoides pubescens*), white-breasted nuthatches, tufted titmice, Carolina chickadees, and house sparrows) tested in both winter and spring and a sixth species (European starlings, *Sturnus vulgaris*) tested only in winter. The clicks were presented at several rates and amplitudes. The authors were surprised to find relatively small differences in the overall average latency and amplitude of the ABR across the six species they studied (i.e., no or small main effect of species in their models). However, the addition of season and the season × species interaction revealed some unexpected patterns. In particular, Lucas et al. (2002) found that the amplitude of the ABR was significantly lower in spring than in winter for nuthatches and woodpeckers, while in chickadees and sparrows, the amplitude of the ABR was greater in spring than in winter.

Lucas et al. (2007) then investigated responses to tones in three species: Carolina chickadees, tufted titmice, and white-breasted nuthatches. They found seasonal plasticity that largely mirrored their previous findings. Chickadees tended to have stronger responses (greater amplitude) in the spring compared to the winter, while nuthatches had greater amplitude responses in the winter compared to the spring. The changes in the chickadees seemed to affect responses across a range of frequencies, while in nuthatches the plasticity was restricted to a narrow range of frequencies (1–2 kHz). Titmice did show seasonal plasticity in their onset response to tones (ABR), but did not show plasticity in their sustained responses (FFRs) .

Henry and Lucas (2009) next investigated frequency sensitivity in the house sparrow. They found that while the amplitude of ABRs to tone burst stimuli showed seasonal plasticity, there was no variation in ABR-based absolute thresholds (the audiogram) or the latency of the ABR response. The amplitudes of the response to tones from 3.2 to 6.4 kHz were greater in the spring than in the fall (Fig. 2.6). Again, these results mirror the initial findings of Lucas et al. (2002), with amplitude of the responses increasing during the spring months, when house sparrows are in reproductive condition.

Gall et al. (2013) and Vélez et al. (2015b) extended the work on seasonality by investigating temporal and frequency resolution during autumn (nonbreeding) and spring (breeding). In the house sparrow, Gall et al. (2013) found that there were no (or very limited) main effects of season on temporal resolution and frequency resolution. However, temporal and frequency resolution did show significant

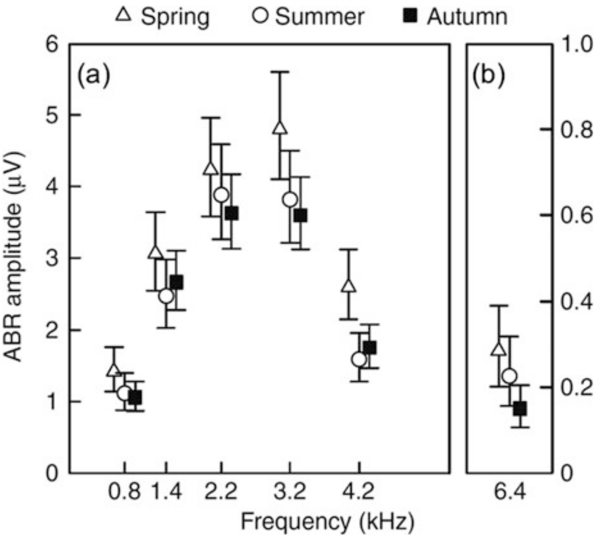


Fig. 2.6 Seasonal variation in the amplitude of tone-evoked ABRs in the house sparrow (from Henry and Lucas 2009 [Fig. 3]). Average ABR amplitude in spring, summer, and autumn (see legend) at frequencies of (**a**) 0.8–4.2 kHz and (**b**) 6.4 kHz. *Error bars* represent 95 % confidence intervals. Note that spring data are offset −150 Hz, autumn data are offset +150 Hz, and amplitude scales differ between panels

seasonal variation when sex and season were considered in concert. Similarly, in other bird species, Vélez et al. (2015b) found that the main effect of season did not influence the response of Carolina chickadees, tufted titmice, or white-breasted nuthatches to the amplitude envelope or the fine structure of amplitude-modulated tones. However, they found that the response of titmice and chickadees to the fine structure of sound (FFR) did differ seasonally when sex and season were considered simultaneously. These sex × season interactions are discussed in greater detail in Sect. 2.6. Finally, Vélez et al. (2015b) found that there were seasonal changes in the EFR in each of the species, but only at a subset of amplitude modulation rates.

The work described above strongly suggests that there is seasonal plasticity in the adult auditory periphery in songbirds. There are two natural questions that follow from these observations: (1) What mechanisms underlie this plasticity? (2) What, if anything, is the function of this plasticity? While there has been some work addressing the first question, at this point we can only speculate about the potential functions of this plasticity.

There are several possible mechanisms that could be responsible for plasticity in the auditory periphery of songbirds. The most likely candidate seems to be that seasonally or reproductively related changes in hormone levels influence the electrical tuning of hair cells. Work in midshipman fish has shown that seasonal changes in auditory sensitivity are correlated with plasma steroid hormone level (Rohmann and Bass 2011). The tuning of hair cells in these fish is largely determined by large conductance ion channels, and seasonal changes in auditory tuning are linked to the expression of big potassium (BK) channels (Rohmann et al. 2009, 2013, 2014). Furthermore, manipulation of BK channels in larval zebra fish leads to changes in auditory sensitivity (Rohmann et al. 2014). In songbirds, estrogen receptors and aromatase have been found in hair cells (Noirot et al. 2009), and exogenous steroid hormones lead to plasticity in the amplitude of the ABR response (Caras et al. 2010). Other possible mechanisms of plasticity, which remain unexplored in songbirds, include the addition of hair cells in the inner ear [hair cell addition has been linked to seasonal changes in sensitivity in fish (Coffin et al. 2012)] or replacement or remodeling of hair cell structure. Songbirds are capable of regenerating hair cells following injury (Marean et al. 1998); however, it is not clear whether this mechanism could be involved in seasonal plasticity. Finally, auditory feedback from mate attraction signals may result in developmental changes that serve to enhance peripheral responses to those calls. This phenomenon has been seen in the auditory periphery of frogs (Gall and Wilczynski 2015) and in higher-order auditory processing areas in songbirds (Sockman et al. 2002, 2005) and frogs (Gall and Wilczynski 2014).

There are several hypotheses regarding the function of seasonal plasticity. Most commonly, this plasticity is assumed to enhance the ability of individuals to respond to vocal signals during the breeding season. However, this could be equally accomplished by having highly sensitive hearing year-round. There are two main mechanisms, then, by which plasticity may be favored over year-round sensitivity. The first is that plasticity may reduce energetic expenditure on sensory tissue during the nonbreeding season, as sensory tissues are expensive to maintain (Niven and Laughlin 2008). A similar mechanism is thought to drive plasticity in neural tissues controlling song production in male birds. An alternative explanation, particularly for plasticity in temporal and frequency resolution, is that these changes allow enhanced detection of seasonally specific sounds in both seasons. For instance, enhanced frequency resolution may be beneficial for females selecting mates on the basis of their vocalizations, while enhanced temporal resolution may be beneficial for both sexes in the nonbreeding season when signals about food or predation risk may be particularly important.

A second hypothesis about the function of plasticity is that plasticity may gate the salience of particular signals, rather than enhance their detection or discrimination. This idea has been largely unexplored in songbirds, although there is some support from other taxa. For instance, female frogs for which egg laying is imminent will respond with positive phonotaxis to male calls that are less attractive, while females earlier in their reproductive trajectory are more discriminating (Lynch et al. 2005, 2006). These changes in behavior are correlated with hormone profiles (Lynch and Wilczynski 2005), which are in turn correlated with changes in midbrain sensitivity to male calls (Lynch and Wilczynski 2008). When reproductive hormone levels are increased by administering gonadotropin, the overall activity of the midbrain is increased, suggesting that even relatively poor calls are likely to activate sensory processing areas when females are close to laying eggs. Therefore, plasticity in the responsiveness of the auditory midbrain to vocalizations may gate the salience of these signals for evoking behavioral responses. Similarly, birds' enhanced sensitivity, or an altered balance of frequency and temporal resolution, may gate the salience of male reproductive signals. This sensory gating could then work in concert with hormonally induced changes in female motivation to modulate reproductive behavior.

2.6 Sex Differences in Auditory Processing

In birds, and in particular songbirds, there are often great differences between males and females in the production of vocalizations. Yet despite the well-studied variation between the sexes in terms of signal production, relatively little work has focused on variation in sensory processing between the sexes. In the auditory system, this oversight seems to have two primary causes. The first is that variation (both among seasons and between the sexes) in behavioral responses to signals has generally been assumed to be due to motivational, rather than due to sensory processing differences. The second is that the most common methodologies used to investigate auditory processing in birds were not well suited to the investigation of sex differences.

Prior to the early 2000s, most of what we knew about general aspects of avian audition was the result of either psychophysics or single-unit electrophysiological recordings. In both circumstances, the number of subjects is necessarily limited and the effect of sex on audition was, therefore, generally not considered. The early work on AEPs in birds was conducted in chickens and ducks, and focused on validating the methodology, on the development of auditory sensitivity, or on broad taxonomic comparisons, but did not investigate sex differences (Saunders et al. 1973; Aleksandrov and Dmitrieva 1992; Dmitrieva and Gottlieb 1994).

Much of the work on AEPs in songbirds has either not considered the sex of the animals (Brittan-Powell and Dooling 2002, 2004; Lucas et al. 2002) or did not find strong effects of sex on auditory responses (Lucas et al. 2007, 2015; Henry and Lucas 2008, 2010b; Caras et al. 2010; Vélez et al. 2015a; Wong and Gall 2015). For instance, Lucas et al. (2007) found no main effects of sex and marginal effects of sex × season and sex × species on the processing of tones when investigating three species of birds, although the within-season sample size for each sex was quite low (one to three individuals). Similarly, Henry and Lucas (2008) found no effects of sex on frequency sensitivity or temporal resolution in three species (tufted titmice, white-breasted nuthatches, and house sparrows) that were tested primarily during the nonbreeding season. Nor did Lucas et al. (2015) find sex differences in the processing of complex tones in four species of songbird sampled during the nonbreeding season. Additionally, when samples are pooled across the breeding and nonbreeding seasons, sex effects have rarely been observed (Henry and Lucas 2010b).

Sex differences were first observed by Henry and Lucas (2009) in an investigation of seasonal patterns of frequency sensitivity in the house sparrow. They found that the ABR amplitude of male house sparrows increased at a greater rate than the ABR amplitude of females as the amplitude of the stimulus increased. They did not find any effects of sex on auditory thresholds or the latency of the ABR response. (They did not investigate the sex × season interaction due to inadequate sample sizes.) Henry and Lucas (2010a) then investigated frequency sensitivity in Carolina chickadees following the breeding season (September to November) but during a time of pair formation (Mostrom et al. 2002). Here they found marginal main effects of sex on ABR thresholds, but a clear effect of the sex × frequency interaction on ABR amplitude and latency. Generally, males tended to have lower thresholds and greater amplitudes, but longer latencies. Additionally, Henry and Lucas (2010a) found that males had greater frequency resolution than females during this time. In the same year, Gall and Lucas (2010) found that frequency resolution of brown-headed cowbirds sampled during the breeding season also showed large differences between the sexes, with females having greater frequency resolution than males. Similarly, in both brown-headed cowbirds and red-winged blackbirds tested during the breeding season, females had slightly lower ABR thresholds and much greater ABR amplitudes than males (Gall et al. 2011).

These early data suggested that sex differences in auditory processing occur but that the exact nature of these differences is both species specific and time specific. In particular, this early work seemed to suggest that sex differences are greatest during times of pair formation and breeding. Additionally, it seemed that these differences were greatest in species with a high degree of sexual dimorphism and that sex differences were greatest in species in which mate assessment occurs over a relatively shorter time period.

The Carolina chickadee and blackbird data led to the hypothesis that sex and season interact to influence auditory processing in songbirds in a way that is similar to how sex and season interact to influence the production of mate attraction signals. In temperate songbirds the utility of mate attraction signals decreases during the nonbreeding seasons; thus, at these times it is more energy efficient to downregulate neural tissue devoted to song production, rather than to maintain these tissues for long periods of disuse. Similarly, auditory processing may be modulated over the course of the year to match the auditory stimuli of greatest importance. Based on this hypothesis, we would predict that the sexes would differ in their auditory processing when females are evaluating mate attraction signals, but that auditory processing of the sexes would converge during the nonbreeding season. These changes could either improve female discrimination of vocalizations or alter the downstream neural population responding to a particular stimulus, thus altering the salience of mate attraction signals during pair formation. In this way, peripheral changes in audition could gate the release of mating behavior, potentially independently or in concert with changes in motivation. This assumes that for males the benefit of enhanced auditory processing during the breeding seasons may not offset the costs or that males and females may process different elements of male vocalizations. During the nonbreeding season, when males and females may need to respond to similar auditory stimuli, such as those related to predation and foraging, auditory processing is expected to converge.

This hypothesis was tested for the first time in house sparrows, with an investigation of frequency and temporal resolution (Gall et al. 2013). Henry et al. (2011) had previously shown that frequency resolution and temporal resolution are inversely related to one another, both at the species and individual level in songbirds. Gall et al. (2013) found that in the nonbreeding season, there was no difference between males and females in temporal resolution or frequency

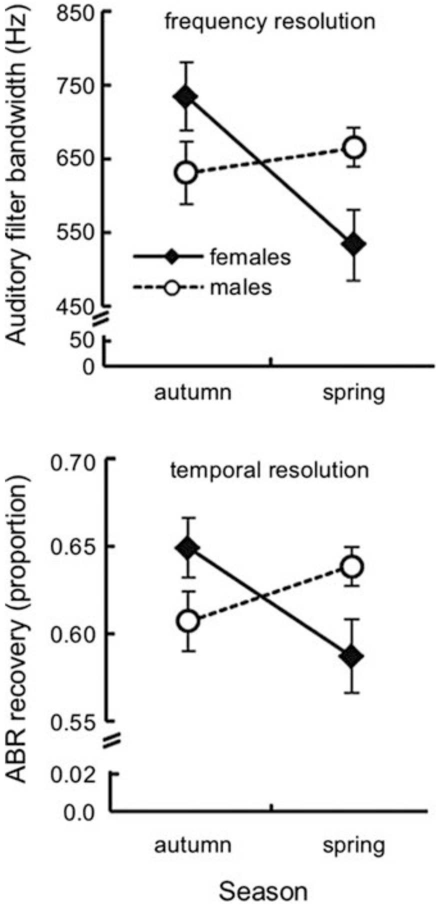


Fig. 2.7 Seasonal variation in ABR-based estimates of frequency resolution (*top*) and temporal resolution (*bottom*) in male and female house sparrows (from Gall et al. 2013 [Fig. 1]). Note that females (*black diamonds*) show an increase in frequency resolution (manifest as a reduction in auditory filter bandwidth) from the nonbreeding season (autumn) to the breeding season (spring) and, as a possible consequence, a concomitant decrease in temporal resolution (reduction in ABR recovery in response to a second click)

resolution. However, during the breeding season, females had greater frequency resolution than males, but poorer temporal resolution. This seasonal pattern seems to be due to plasticity in females, as males did not exhibit auditory variation between the seasons, while females did (Fig. 2.7).

This pattern of sex-specific seasonal plasticity in frequency resolution and temporal resolution, as well as in the processing of pure tones, has now been observed in many other species. For instance, Vélez et al. (2015b) found that in the winter there were no sex differences in the EFR, a measure of temporal resolution, in three species of songbirds (tufted titmice, white-breasted nuthatches, and Carolina chickadees). However, in the spring there were significant differences between the sexes in their ability to follow certain rates of temporal modulation. In titmice and in chickadees, males tended to have a greater ability than females to follow temporal modulation in the spring. Similarly, the strength of FFRs to pure tones did not differ between the sexes during the winter, but did differ during the spring. Typically, females had greater FFR amplitude in response to pure tones than males during the spring, with no difference in the winter.

The data from titmice, white-breasted nuthatches, and Carolina chickadees tell a similar story to that of the house sparrows and red-winged blackbirds: frequency resolution improves in females during the breeding season at the expense of temporal processing. Frequency sensitivity, when measured as the amplitude of the ABR to the onset of single frequency sound stimulus, also tends to be greater in females than in males during the breeding season (Gall et al. 2012b). However, other types of auditory processing tend to be less plastic. For instance, in these same species, there are only small differences in ABR-based auditory thresholds, even during the breeding season (Vélez et al. 2015a). Therefore, different facets of auditory processing in songbirds should be treated independently when investigating sex effects. Future research should strongly consider both season and stimulus type when investigating sex-specific auditory processing.

2.7 Individual Variation in Auditory Physiology

There is enormous variation at several scales in the structure of signals: among species (Gerhardt and Huber 2002; Catchpole and Slater 2008), among individuals of the same species (e.g., between sexes; Catchpole and Slater 2008), and within individuals over time (Hill et al. 2015; Maddison et al. 2012). This variation in signal structure, or the phonemes or other elements characteristic of a signal, is often related to the function of the signal (Bradbury and Vehrencamp 2011). Signal function also influences, in part, the level of variation in the structure of signals at the species and individual levels. For example, signals used in interspecies communication tend to converge at the species level: mobbing or alarm calls share similar acoustic features between species that flock together (Ficken and Popp 1996; Hurd 1996). Interestingly, these calls also tend to be similar at the individual level. In contrast, elements of signals that denote species identity (i.e., species badges) are known to diverge among species, but individuals of the same species tend to share these signal elements (Gerhardt 1991). Finally, honest signals that denote individual quality diverge in structure at the species level (i.e., different species use different signals), and individuals within a species also differ (Grafen 1990; Searcy and Nowicki 2005).

In this chapter we have discussed a parallel set of scales with respect to variation in the auditory system: among species, between sexes, and within individuals over time (also see Dangles et al. 2009). The functional attributes of this variation tend to be fairly straightforward: species differences in auditory processing reflect, in part, constraints imposed by habitat on signal propagation (Morton 1975; Wiley 1991); sex differences in auditory processing may reflect differences in the requirement for processing sex-specific aspects of vocal signals (Gerhardt and Huber 2002); variation within individuals over time may reflect several factors such as seasonal variation in the use of different signal types (Catchpole and Slater 2008).

One of the smallest ecological scales we can consider is variation within-sex and within-season. Here the correspondence between signal evolution and the evolution of sensory systems seems to be relatively under-explored (Ronald et al. 2012). Indeed, Dangles et al. (2009) discussed variability in sensory ecology across a variety of scales, from populations to individual development, but they left out any discussion of variability between individuals of the same age and sex.

While the relationship within individuals between signal evolution and sensory evolution is poorly understood, the functional aspects of individual variability in signal design itself are well documented. Indeed, the signal side of individual variability is the heart of our theoretical framework for sexual selection. For example, the classic handicap principle offers a hypothesis for why males should differ in the intensity of a signal (Zahavi 1975; Grafen 1990): signals will evolve to be costly to the signaler if the magnitude of the cost is relatively higher for low-quality signalers compared to high-quality signalers. However, this theory is based on the implicit assumption that a signal of a given intensity, and therefore the information contained within that signal, is a fixed entity with a characteristic cost (e.g., Grafen 1990; Johnstone 1995; Searcy and Nowicki 2005). Logically, if the value of a signal with a specific set of properties is to be treated as a fixed entity, then either all receivers will have to process that signal in an identical way or the mapping of signal properties on to signal information content (see Bradbury and Vehrencamp 2011; Hailman 2008) has to result in an identical signal valuation across receivers. We have known for some time that receivers do vary in their response to signals, although the factors that have been addressed are primarily factors that affect functional aspects of mating decisions. For example, individual variation in mate choice is affected by a variety of factors, such as the physical and social signaling environment (Gordon and Uetz 2011; Clark et al. 2012), previous experience (Bailey 2011; Wong et al. 2011), genetic differences between choosing individuals (Chenoweth and Blows 2006; Horth 2007), and female condition (Cotton et al. 2006).

The treatment of the signal as a fixed entity ignores potential variability in the sensory processing capabilities of receivers. Additionally, if signal processing varies among individuals (i.e., individuals differ in their capacity to extract information from a signal), then signal information content decoded from the signal may also vary among individuals, particularly in the case of complex signals (Kidd et al. 2007). This individual variation in sensory processing is important theoretically because it has the potential for altering the fitness consequences of signals in several ways. First, the relative fitness accrued from expressing a particular signal will become more variable if signals are detected and processed differently by different receivers. Second, if individual variation in sensory processing varies among groups of receivers (e.g., age groups or groups of individuals that vary in their exposure to various sounds), then signals will vary in efficacy depending on the specific group to which they are directed. This raises the question as to whether signal processing does, in fact, vary among individuals and whether this variation is group specific. The answer to both of these questions is “yes.”

Individual variation in signal processing that is likely to alter mate-choice decisions has been demonstrated in several systems. Henry et al. (2011) found individual variation in ABR-based auditory filter bandwidth in a sample of Carolina chickadees. Moreover, they demonstrated that chickadees with broader filters had greater auditory temporal resolution (measured with paired-click stimuli) than chickadees with narrow filters. Thus, individual chickadees vary in the degree to which they are able to resolve temporal cues (such as amplitude modulation) and spectral cues (such as frequency properties) in any given vocal signal. Similarly, Ensminger and Fernandez-Juricic (2014) found individual variation in cone density in the eyes of house sparrows. Chromatic contrast models were used to illustrate that these differences in cone density would result in differences in the ability of females to detect the quality of plumage signals known to be associated with mate choice. This individual variation in signal processing may alter the capacity of the signaler to encode relevant information in a signal, and it may also limit the capacity of the receiver to decode that information.

Group-specific variation is most easily shown in animals that are subject to different environments. For example, Gall and Wilczynski (2015) demonstrated that green treefrogs (*Hyla cinerea*) that were exposed to species-specific vocal signals as adults had altered peripheral auditory sensitivity compared to frogs that were not exposed to vocal signals. Phillmore et al. (2003) found that black-capped chickadees (*Poecile atricapillus*) that are isolate-reared had more difficulty in identifying individual-specific vocalizations than field-reared birds. These results are similar to those of Njegovan and Weisman (1997), who showed that isolate-reared chickadees also have impaired pitch discrimination.

There are three potential mechanisms that generate individual variation in auditory physiology: age-related differences, experience-related differences, and hormone-related differences. All of these mechanisms can be the cause of striking phenotypic plasticity in the auditory system. Age-related effects on hearing are well documented in humans (Clinard et al. 2010; He et al. 2007; Mills et al. 2006; Pichora-Fuller and Souza 2003) and in a variety of model systems (e.g., mice, *Mus musculus*, Henry 2002; rats, *Rattus norvegicus*, Parthasarathy and Bartlett 2011). Differences in rearing conditions potentially shape many aspects of auditory processing, from frequency coding and tonotopic maps to spatial processing and vocalization coding (Sanes and Woolley 2011; Woolley 2012; Dmitrieva and Gottlieb 1994). The auditory system is relatively plastic, even in adults. AEPs can change in frogs as a result of exposure to simulated choruses (Gall and Wilczynski 2015). Patterns of AEPs can be altered in a way that indicates better auditory processing in people with hearing deficits who undergo auditory training (Russo et al. 2005). Musicians compared to nonmusicians show enhanced auditory processing of tones (Musacchia et al. 2007), as do people who speak tonal languages such as Mandarin (Krishnan and Gandour 2009). Musicians are better at solving the cocktail party problem (i.e., detecting and recognizing speech in noise; see Miller and Bee 2012) than are nonmusicians (Swaminathan et al. 2015). These examples show an explicit link between experience-dependent auditory plasticity and vocal communication.

Hormone levels, and particularly estrogen levels, are known to affect both developmental and activational levels of auditory performance (Caras 2013). Hormone effects are particularly important in seasonal plasticity of the auditory system where they promote retuning of the auditory system in a wide variety of taxa, including fish (Sisneros 2009), anurans (Goense and Feng 2005), birds (Caras et al. 2010), and mammals (Hultcrantz et al. 2006). However, as discussed earlier in this chapter, the details of this retuning can be quite complex. These hormone effects can, in turn, influence individual variability if there are individual differences in either the timing of the reproductive cycle or in the mean amplitude of estrogen levels during that cycle.

Charles Darwin's theory of natural selection (Darwin 1872) is the framework within which almost all of biological thought is built. One of the three tenets of this theory is variation between individuals within a population. It seems almost odd that we spend so little time thinking about this in the context of sensory processing.

2.8 AEP Responses to Natural Vocalizations

In this section we describe some new analyses of AEPs in response to natural vocalizations that illustrate a number of factors we have discussed in this chapter. Virtually all of our own work has focused on the auditory processing of either simple sounds (clicks and tones) or sounds that mimic elements of vocal signals, such as harmonic tone complexes and amplitude-modulated sounds. While these studies provide critical information about auditory processing, it is particularly important to address how auditory systems process real vocal signals. This approach has led to important insights into hearing deficits in humans (Johnson et al. 2005, 2008) and into the processing of songs in the avian auditory forebrain (Amin et al. 2013; Elie and Theunissen 2015; Lehongre and Del Negro 2011).

The beauty of AEPs is that they can be generated with any input stimulus, including natural vocal signals. The processing of a vocal signal will generate a complex AEP waveform that can be quantified in a number of ways, although we take only a single approach here: cross correlation of response waveforms. We restrict our analysis to three signals: a tufted titmouse (hereafter ETTI) song element, a white-breasted nuthatch (hereafter WBNU) song element, and a WBNU contact ("quank") call element (Fig. 2.8). The WBNU song and quank elements are interesting because their spectrograms are similar but their function is different: the song is used in mating-related contexts and the quank is both a contact call and a call used in mobbing. The ETTI song is structurally different than the WBNU song but shares the same function. Two exemplars of each signal were used in the experiments and are illustrated in Fig. 2.8; AEPs generated by these signals are shown in Fig. 2.9. Note that both of these figures depict two views of a signal (Fig. 2.8) or the AEP response (Fig. 2.9): the top figures (waveform plots) depict the amplitude of the pressure waveform or voltage waveform, respectively, as a function of time. The bottom figures (spectrogram plots) are a Fourier transform of

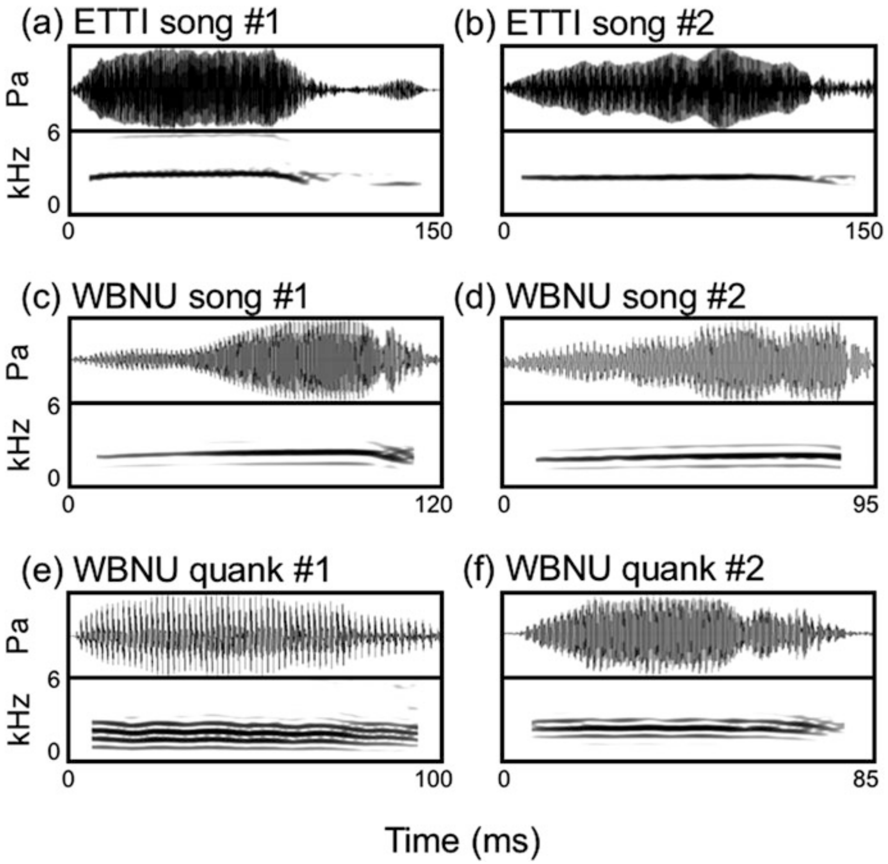


Fig. 2.8 Natural vocalizations used to elicit AEPs. Shown here are the six input stimuli used in our study of AEPs in response to natural vocalizations of the tufted titmouse (ETTI) and white-breasted nuthatch (WBNU): (**a**) ETTI song #1, (**b**) ETTI song #2, (**c**) WBNU song #1, (**d**) WBNU song #2, (**e**) WBNU quank #1, (**f**) WBNU quank #2. The waveform (*top*) and spectrogram (*bottom*) are shown for each vocalization. The y-axis labels are identical for all figure pairs

the corresponding waveform, and they depict the frequency content of the signal (Fig. 2.8) or the AEP response (Fig. 2.9) as a function of time. In the spectrogram, simultaneous bands represent signal elements (Fig. 2.8) or evoked potential elements (Fig. 2.9) that include multiple frequency components. The technique used to measure AEPs, summarized briefly earlier in this chapter, is described in detail in Vélez et al. (2015a), Lucas et al. (2015), Gall et al. (2013), and Henry et al. (2011); we will not cover the technique here.

The two WBNU songs are composed of several harmonics that generate a fairly strong amplitude modulation, or beating, at a rate equal to the difference in frequency between the harmonics (Table 2.1). Here is a simple example to illustrate this: say you have three simultaneous tones with frequencies of 1200, 1800, and

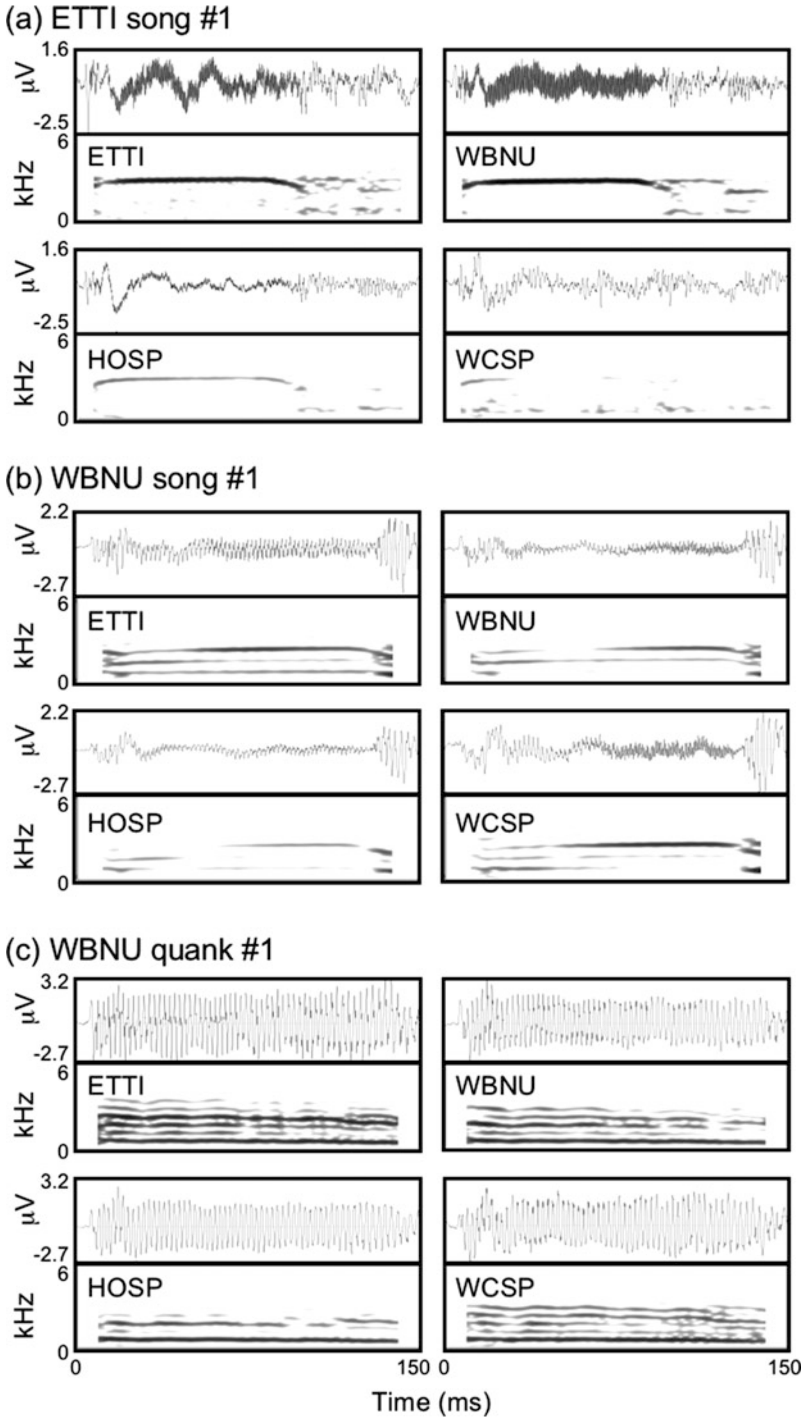


Fig. 2.9 AEPs in response to natural vocalizations. Shown here are mean AEP responses to one exemplar of each of the three types of stimuli played to each of the four species tested. Each figure

2400 Hz. These tones represent a harmonic stack with a lower fundamental frequency of 600 Hz, although this fundamental is missing in our example. However, the three tones combine in a way that generates an amplitude modulation of 600 Hz. The WBNU songs and calls have this exact structure. This is important, because the auditory system will phase-lock to that AM rate in addition to phase-locking to each separate tone (Henry and Lucas 2008; Lucas et al. 2015; Vélez et al. 2015b). In addition to this AM component, all of these WBNU songs also have a gross amplitude envelope that starts at a low intensity in the beginning of the song element and has a peak intensity toward the end of the element (Fig. 2.8c, d).

Like the WBNU song, the quank call notes (Fig. 2.8e, f) can be characterized as having a series of strong harmonics that generate strong AM in both call notes (Table 2.1). What differs between the WBNU quank and WBNU song elements is that the WBNU song has a higher AM rate than the quank calls (Table 2.1), and the gross amplitude envelope increases more slowly (compare Fig. 2.8c, d vs. Fig. 2.8e, f). Both titmouse song elements are relatively tonal, with fundamental frequencies of about 2.6 kHz (Table 2.1; Fig. 2.8a, b). However, the song elements differ in their gross amplitude envelopes. Amplitude onset is slower in the ETTI song element 2 (Fig. 2.8b) than in the ETTI song element 1 (Fig. 2.8a). ETTI song element 1 also has a weak harmonic at 5.4 kHz that is not evident in the spectrogram of ETTI song element 2.

We measured AEPs from four species of birds: two tufted titmice (one male, one female), six white-breasted nuthatches (three males, three females), six house sparrows (five males, one female), and six white-crowned sparrows (three males, three females). Each bird was tested on both exemplars of each of the three signals. The average evoked potentials in response to the ETTI song element, WBNU song element, and WBNU quank call are given in Fig. 2.9.

We can use cross correlation (estimated with the Praat program; Boersma and Weenink 2009) to compare the AEP waveforms resulting from each stimulus measured in all pairs of birds in our sample, including both intraspecific pairs and interspecific pairs. Any two waveforms will show a cross correlation of nearly 1.0 if they are nearly identical; if they are quite different, they will show a cross correlation of nearly 0.0. The cross correlation analysis yields a matrix of correlations between all pairs of waveforms. We reduced the dimensionality of the matrix using multidimensional scaling (MDS; Proc MDS, SAS Institute Inc., v9.1). MDS is similar to principal component analysis in the sense that it estimates the relative position of a set of objects (auditory responses of individual birds in this case) in a space of a specified number of dimensions. Our MDS analysis used absolute values of dissimilarity (i.e., 1 minus the cross correlation coefficient), which represents

◀

Fig. 2.9 (continued) has two parts: the *top* is the waveform view and the *bottom* is the spectrogram derived from the waveform. AEPs are shown for the tufted titmouse (ETTI), white-breasted nuthatch (WBNU), house sparrow (HOSP), and white-crowned sparrow (WCSP) in response to (**a**) ETTI song #1, (**b**) WBNU song #1, and (**c**) WBNU quank #1. Note: the voltage intensity (*y*-axis) is the same for waveforms resulting from the same input stimulus

Table 2.1 Characteristics of exemplars used to generate AEPs to natural vocal signals

	ETTI song		WBNU song		WBNU quank	
Property	#1	#2	#1	#2	#1	#2
Harmonics (kHz)	2.8	2.5	1.6	1.4	1.1	1.7
	5.4		2.4	2.1	1.6	2.3
			3.2	2.8	2.2	2.8
					2.7	
AM rate of maximum intensity (Hz)	2800		774	693	546	569
Frequency of maximum intensity (kHz)	2.8	2.5	2.4	2.1	2.2	2.3

Two exemplars (#1 and #2) were used for each stimulus type. Harmonics, AM rate of maximal intensity, and frequency of maximal intensity were determined using a spectrum calculated with Praat software (Boersma and Weenink 2009)

approximate levels of dissimilarity between AEP waveforms of each bird across MDS space. The MDS procedure generates a badness-of-fit statistic that can be used to estimate an approximate R (badness-of-fit $= \sqrt{1 - R^2}$, where R is the multiple correlation about the origin). For our analyses, $R > 0.95$ for all MDS models using two dimensions.

Our results, summarized in Fig. 2.10, illustrate three patterns about variability in auditory physiology. First, birds from the same habitat (woodland habitat, titmice and nuthatches; open habitat, sparrows) tend to be more similar to each other than to the birds from a different habitat (species effect in a MANOVA: both ETTI songs, $F_{9,22} > 5.0$, $P < 0.001$; both WBNU songs, $F_{9,56} > 7.0$, $P < 0.001$; both WBNU quanks, $F_{9,22} > 3.0$, $P < 0.024$). Second, the species are much more similar to each other in response to the quank call compared to their response to either of the songs (ANOVA of between-species standard deviation estimates, $F_{2,3} = 143.2$, $P = 0.001$). Finally, individual variation is significantly greater in response to both songs compared to the response to the quank call (ANOVA of within-species standard deviation estimates, $F_{2,18} = 35.2$, $P < 0.001$; Fig. 2.11). Thus, variability among species mirrors variability within species.

Why should the variability be so different among signals? These three signals can be characterized using three predominant properties. (1) Maximal energy for each of them is located at about 2–3 kHz. (2) The amplitude modulation rates are about 725 Hz for WBNU song and about 550 Hz for the WBNU quank calls. Titmice have a weak amplitude modulation rate at about 2.5 kHz. (3) With respect to the gross amplitude envelope, the rate of amplitude increase at the beginning of the call, which is higher in the quank call than in the songs, may also be important.

Our previous results show that the AM rate that generates the strongest auditory response across a range of species is about 400–600 Hz (Henry and Lucas 2008; Gall et al. 2011). Similarly, 2–3 kHz is the range of most sensitive hearing for many songbirds (Dooling et al. 2000). Thus, the quank call appears to be near the best AM rate and frequency for optimal processing for many species. These properties may be the basis for the low variation in AEPs both within and between species. The rapid onset of this signal may also aid in auditory processing.

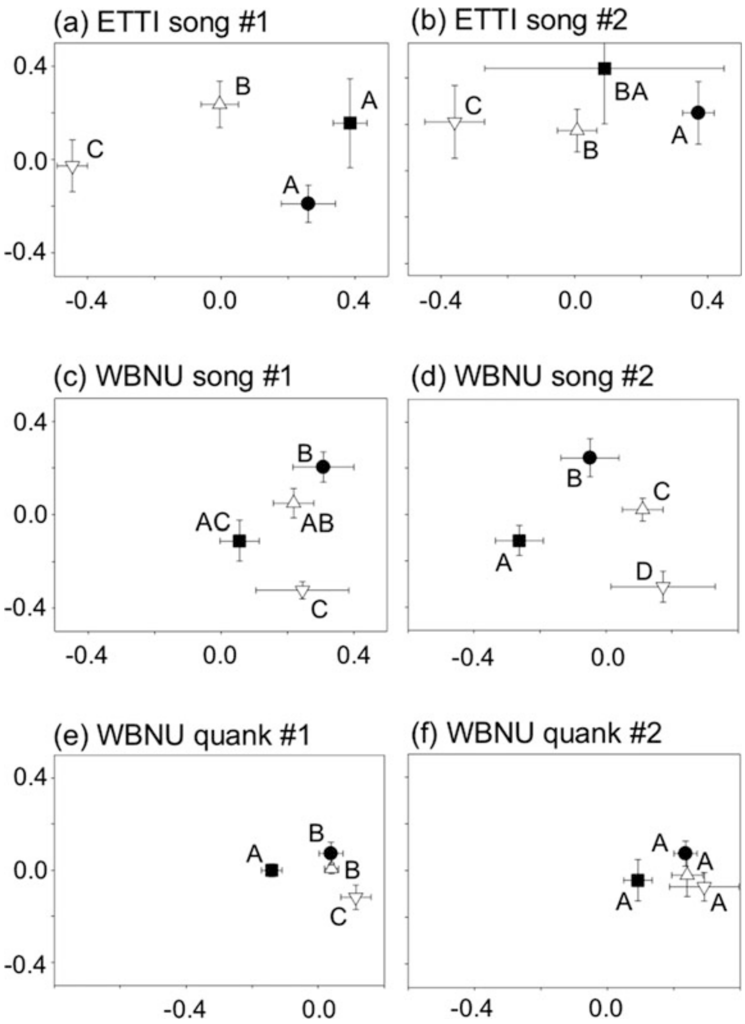
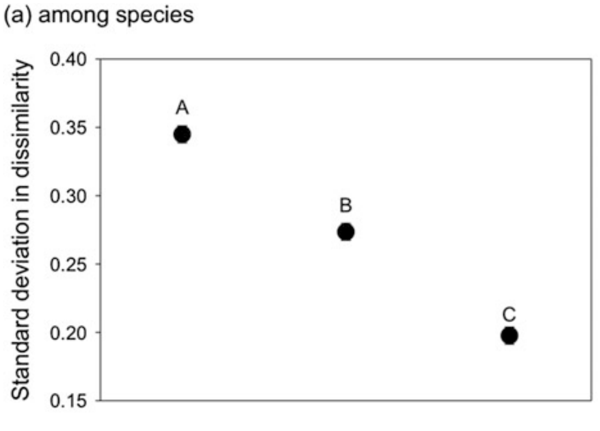


Fig. 2.10 Analyses of species differences and habitat-related differences in AEPs elicited by natural vocalizations. Shown here are mean ± s.d. dissimilarity values for the tufted titmouse (ETTI), white-breasted nuthatch (WBNU), house sparrow (HOSP), and white-crowned sparrow (WCSP) based on multidimensional scaling analysis for AEP waveforms in response to (**a**) ETTI song #1, (**b**) ETTI song #2, (**c**) WBNU song #1 (**d**) WBNU song #2, (**e**) WBNU quank #1, and (**f**) WBNU quank #2. *Filled symbols* indicate species living in woodland habitat (*filled square* ETTI and *filled circle* WBNU). *Open symbols* indicate species living in open habitat (*open up-pointing triangle* HOSP and *open down-pointing triangle* WCSP). Symbols with the same letter are not significantly different from one another based on a MANOVA

The fact that the level of variance is higher for song elements than it is for the quank call is intriguing because the quank call is used as a mobbing call that is designed to draw in a variety of different species (Ficken and Popp 1996; Hurd



(b) within species

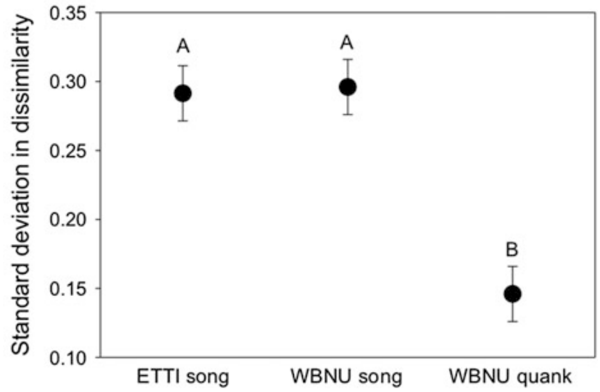


Fig. 2.11 Standard deviation in dissimilarity values from a multidimensional scaling analysis. The dissimilarity values were derived from pair-wise cross correlations between the AEP waveforms to natural sounds for all pairs of birds (among species and within species) in our study. These standard deviation values are for the dissimilarity estimates represented in Fig. 2.10. (**a**) Variation in pair-wise comparisons among species; (**b**) variation in pair-wise comparisons within species. Values are least squares means ± s.e.m. from ANOVA models (see text). Symbols with the same letter are not significantly different from one another based on an ANOVA

1996). The low variance within species should increase the probability that each receiver interprets the signal correctly. Moreover, the distance over which mobbing calls can be detected (see Lohr et al. 2003) should be maximized if the mobbing calls are designed with peak energy at 2–3 kHz and AM rates of about 400–600 Hz.

The high interspecific variability in response to the song elements is not surprising given that the songs should be designed for their species-specific targets. What is quite intriguing is that this variability is mirrored intraspecifically. This could alter the nature of sexual selection in these species, in part because it weakens selection on the signal by potentially increasing the variance in receiver decoding (as described in Sect. 2.7). Moreover, our results suggest that signalers may design signals that differentiate between receivers through differences in receiver sensory capabilities.

2.9 Summary and Future Directions

There are several themes to this chapter. The main theme is an old one: we can evaluate auditory performance at a host of hierarchical levels, and information from each level enhances our understanding of how animals adapt to their sensory environment. What is clear from our work on songbirds is that a crucially important component of that environment is vocal signaling. Indeed, vocal signal properties (and particularly the properties of species-specific song) seem to be a major driver of auditory performance in all of the species we have studied. How the feedback loop between signal/signaler properties and receiver perceptual properties evolves is an extremely important but extremely difficult question to answer.

The second theme of our chapter underscores this point: the auditory system is incredibly plastic at every hierarchical level we have discussed here. The details of this plasticity are quite complicated. We currently do not know much about the mechanistic basis of this plasticity, and, therefore, we cannot understand at a particularly deep level the evolutionary trajectories of the auditory properties discussed in this chapter. This is why we have generally refrained from talking explicitly about evolution per se.

The final theme of this chapter is technical: auditory evoked potentials have proven to be an extraordinarily efficient way to characterize the subcortical processing of sound, at least in our study species. With AEPs, it has been possible to characterize one or more of a host of auditory properties in an individual bird (e.g., auditory filter bandwidth, FFRs, EFRs, the audiogram, processing of a variety of complex, and natural sounds) during a recording session lasting less than 2 h. These experiments would take many hours with single-cell recordings and perhaps weeks (even if the data could be collected) with behavioral measures. Plus, we can test nearly any species we collect in the field and simply release the bird the day after sampling for possible repeated measurements in the future. This is not to say that AEPs tell us everything about the mechanisms of the auditory pathway, nor that they are a perfect measure of auditory physiology. This issue was addressed in the beginning of this chapter. But the technique at the very least gives us a snapshot of auditory performance that offers tremendous insight into how birds gather and process auditory information.

Where do we go from here? Having demonstrated auditory variation in songbirds over a series of hierarchical scales, our approach has primarily been to offer a functional framework to explain these patterns. However, these hypothetical functional relationships have not been tested. Are there behavioral consequences to the seasonal plasticity shown in the auditory system? Are there explicit functional advantages to the nonbreeding auditory states? If so, what are they? The notion that seasonal changes in auditory responsiveness might act as a gating mechanism rather than (or in addition to) functioning to enhance or selectively alter auditory processing of behaviorally relevant sounds is an intriguing idea. Can we find evidence for or against this potential role of auditory plasticity? All of these questions are important and relevant to our understanding of the role that sensory systems play in the ability of animals to adapt to their environments. We have no particularly deep answer to any of them at present.

Another unanswered follow-up question to much of what we covered in this chapter relates to mechanisms. We know with some certainty that estrogen levels play a critical role in the seasonal plasticity of the auditory system, at least in females. However, the details of this seasonal plasticity are species- and even individual-specific. As such, the role of androgens is not just an upregulation of the auditory system, but a very fine-scale alteration of components of that system. Understanding the mechanisms underlying these detailed changes should be a high priority for future research.

Acknowledgments Our studies of the auditory system started as a result of a collaboration between JRL and Todd Freeberg, Ananthanarayan (Ravi) Krishnan, and Glenis Long, all in the lab of Ravi. Ravi was also kind enough to lend us his TDT to continue these studies. Much of this work was funded by an NSF grant (IOS-1121728) to JRL and an NSF doctoral dissertation improvement grant (IOS-1109677) to MG and an Animal Behavior Society graduate student research award to MG.

References

Aleksandrov LI, Dmitrieva LP (1992) Development of auditory sensitivity of altricial birds: absolute thresholds of the generation of evoked potentials. Neurosci Behav Physiol 22:132–137

Amin N, Gastpar M, Theunissen FE (2013) Selective and efficient neural coding of communication signals depends on early acoustic and social environment. PLoS One 8(4):e61417

Bailey NW (2011) Mate choice plasticity in the field cricket *Teleogryllus oceanicus*: effects of social experience in multiple modalities. Behav Ecol Sociobiol 65:2269–2278

Beckers GJL, TenCate C (2001) Perceptual relevance of species-specific differences in acoustic signal structure in Streptopelia doves. Anim Behav 62:511–518

Boersma P, Weenink D (2009) Praat: doing phonetics by computer version 5.1.17 (computer program). http://www.praat.org/. Retrieved 7 May 2009

Boncoraglio G, Saino N (2007) Habitat structure and the evolution of bird song: a meta-analysis of the evidence for the acoustic adaptation hypothesis. Funct Ecol 21:134–142

Bradbury JW, Vehrencamp SL (2011) Principles of animal communication. Sinauer Associates, Sunderland

Brenowitz EA (2004) Plasticity of the adult avian song control system. Ann NY Acad Sci 1016:560–585. doi:10.1196/annals.1298.006

Brittan-Powell EF, Dooling RJ (2004) Development of auditory sensitivity in budgerigars (*Melopsittacus undulatus*). J Acoust Soc Am 115:3092–3102

Brittan-Powell EF, Dooling RJ, Gleich O (2002) Auditory brainstem responses in adult budgerigars (*Melopsittacus undulatus*). J Acoust Soc Am 112:999–1008. doi:10.1121/1.1494807

Calford MB, Webster WR, Semple MM (1983) Measurement of frequency selectivity of single neurons in the central auditory pathway. Hear Res 11:395–401

Capranica RR, Moffat AJM (1983) Neurobehavioral correlates of sound communication in anurans. In: Ewert JP, Capranica RR, Ingle D (eds) Advances in vertebrate neuroethology. Plenum Press, New York, pp. 701–730

Caras ML (2013) Estrogenic modulation of auditory processing: a vertebrate comparison. Front Neuroendocrinol 34:285–299

Caras ML, Brenowitz E, Rubel EW (2010) Peripheral auditory processing changes seasonally in Gambel's white-crowned sparrow. J Comp Physiol A 196:581–599

Catchpole CK, Slater PJB (2008) Bird song: biological themes and variations. Cambridge University Press, Cambridge

Chenoweth SF, Blows MW (2006) Dissecting the complex genetic basis of mate choice. Nature Rev Genet 7:681–692

Christie PJ, Mennill DJ, Ratcliffe LM (2004) Pitch shifts and song structure indicate male quality in the dawn chorus of black-capped chickadees. Behav Ecol Sociobiol 55:341–348. doi:10.1007/s00265-003-0711-3

Clark DL, Roberts AJ, Uetz GW (2012) Eavesdropping and signal matching in visual courtship displays of spiders. Biol Lett 8:375–378

Clinard CG, Tremblay KL, Krishnan AR (2010) Aging alters the perception and physiological representation of frequency: evidence from human frequency-following response recordings. Hear Res 264:48–55

Coffin AB, Mohr RA, Sisneros JA (2012) Saccular-specific hair cell addition correlates with reproductive state-dependent changes in the auditory saccular sensitivity of a vocal fish. J Neurosci 32:1366–1376

Cotton S, Small J, Pomiankowski A (2006) Sexual selection and condition-dependent mate preferences. Curr Biol 16:R755–R765

Dangles O, Irschick D, Chittka L, Casas J (2009) Variability in sensory ecology: expanding the bridge between physiology and evolutionary biology. Q Rev Biol 84(1):51–74

Darwin C (1872) The origin of species by means of natural selection, or the preservation of favoured races in the struggle for life, 6th edn. John Murray, London

Derryberry E (2009) Ecology shapes birdsong evolution: Variation in morphology and habitat explains variation in white-crowned sparrow song. Am Nat 174:24–33

Dmitrieva LP, Gottlieb G (1994) Influence of auditory experience on the development of brainstem auditory evoked potentials in mallard duck embryos and hatchlings. Behav Neural Biol 61:19–28

Dolphin WF, Mountain DC (1992) The envelope following response: scalp potentials elicited in the mongolian gerbil using sinusoidally AM acoustic signals. Hear Res 58:70–78. doi:10.1016/0378-5955(92)90010-K

Dooling RJ (1982) Auditory perception in birds. In: Kroodsma DE, Miller EH (eds) Acoustic communication in birds. Academic, New York, pp. 95–130

Dooling RJ (1992) Hearing in birds. In: Webster DB, Fay RR, Popper AN (eds) The evolutionary biology of hearing. Springer, New York, pp. 545–559

Dooling RJ, Zoloth SR, Baylis JR (1978) Auditory sensitivity, equal loudness, temporal resolving power, and vocalizations in the house finch (*Carpodacus mexicanus*). J Comp Physiol Psychol 92:867–876

Dooling RJ, Lohr B, Dent ML (2000) Hearing in birds and reptiles. In: RJ D, AN P, RR F (eds) Comparative hearing: birds and reptiles. Springer, New York, pp. 308–359

Elie JE, Theunissen FE (2015) Meaning in the avian auditory cortex: neural representation of communication calls. Eur J Neurosci 41(5):546–567

Endler JA (1992) Signals, signal conditions, and the direction of evolution. Am Nat 139:S125–S153

Ensminger AL, Fernandez-Juricic E (2014) Individual variation in cone photoreceptor density in house sparrows: implications for between-individual differences in visual resolution and chromatic contrast. PLoS One 9(11):e111854

Ficken MS, Popp J (1996) A comparative analysis of passerine mobbing calls. Auk 113:370–380

Gall MD, Lucas JR (2010) Sex differences in auditory filters of brown-headed cowbirds (*Molothrus ater*). J Comp Physiol A 196:559–567

Gall MD, Wilczynski W (2014) Prior experience with conspecific signals enhances auditory midbrain responsiveness to conspecific vocalizations. J Exp Biol 217:1977–1982

Gall MD, Wilczynski W (2015) Hearing conspecific vocal signals alters peripheral auditory sensitivity. Proc R Soc B 282:20150749

Gall MD, Brierley LE, Lucas JR (2011) Species and sex effects on auditory processing in brown-headed cowbirds and red-winged blackbirds. Anim Behav 81:973–982

Gall MD, Brierley LE, Lucas JR (2012a) The sender-receiver matching hypothesis: support from the peripheral coding of acoustic features in songbirds. J Exp Biol 215:3742–3751

Gall MD, Henry KS, Lucas JR (2012b) Two measures of temporal resolution in brown-headed cowbirds (*Molothrus ater*). J Comp Physiol A 198:61–68

Gall MD, Ronald K, Bestrom E, Lucas JR (2012c) Effects of habitat and urbanization on the active space of brown-headed cowbird song. J Acoust Soc Am 132:4053–4062

Gall MD, Salameh TS, Lucas JR (2013) Songbird frequency selectivity and temporal resolution vary with sex and season. Proc R Soc B 280:20122296

Gerhardt HC (1991) Female mate choice in treefrogs: static and dynamic acoustic criteria. Anim Behav 42:615–635

Gerhardt HC, Huber F (2002) Acoustic communication in insects and anurans: common problems and diverse solutions. University of Chicago Press, Chicago

Goense JBM, Feng AS (2005) Seasonal changes in frequency tuning and temporal processing in single neurons in the frog auditory midbrain. J Neurobiol 65:22–36

Gordon SD, Uetz GW (2011) Multimodal communication of wolf spiders on different substrates: evidence for behavioural plasticity. Anim Behav 81:367–375

Grafen A (1990) Biological signals as handicaps. J Theor Biol 144:517–546

Greenewalt CH (1968) Bird song: acoustics and physiology. Smithsonian Institution Press, Washington, DC

Hailman JP (2008) Coding and redundancy: man-made and animal-evolved signals. Harvard University Press, Cambridge, MA

Hall JW (1992) Handbook of Auditory Evoked Responses. Allyn and Bacon, Boston

Hartmann WM (1998) Signals, sound, and sensation. AIP Press, Woodbury

He N, Mills JH, Dubno JR (2007) Frequency modulation detection: effects of age, psychophysical method, and modulation waveform. J Acoust Soc Am 122(1):467–477

Henry KR (2002) Sex- and age-related elevation of cochlear nerve envelope response (CNER) and auditory brainstem response (ABR) thresholds in C57BL/6 mice. Hear Res 170:107–115

Henry KS, Lucas JR (2008) Coevolution of auditory sensitivity and temporal resolution with acoustic signal space in three songbirds. Anim Behav 76:1659–1671

Henry KS, Lucas JR (2009) Vocally correlated seasonal auditory variation in the house sparrow (*Passer domesticus*). J Exp Biol 212:3817–3822

Henry KS, Lucas JR (2010a) Auditory sensitivity and the frequency selectivity of auditory filters in the Carolina chickadee, *Poecile carolinensis*. Anim Behav 80:497–507

Henry KS, Lucas JR (2010b) Habitat-related differences in the frequency selectivity of auditory filters in songbirds. Funct Ecol 24:614–624

Henry KS, Gall MD, Bidelman GM, Lucas JR (2011) Songbirds trade off auditory frequency resolution and temporal resolution. J Comp Physiol A 197:351–359

Hill SD, Amiot C, Ludbrook MR, Ji WH (2015) Seasonal variation in the song structure of tui (*Prosthemadera novaeseelandiae*). N Z J Ecol 39(1):110–115

Horth L (2007) Sensory genes and mate choice: evidence that duplications, mutations, and adaptive evolution alter variation in mating cue genes and their receptors. Genomics 90:159–175

Hultcrantz M, Simonoska R, Stenberg AE (2006) Estrogen and hearing: a summary of recent investigations. Acta Otolaryngol 126:10–14

Hurd CR (1996) Interspecific attraction to the mobbing calls of black-capped chickadees (*Parus atricapillus*). Behav Ecol Sociobiol 38:287–292

Johnson DH (1980) The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. J Acoust Soc Am 68:1115–1122

Johnson KL, Nicol TG, Kraus N (2005) Brain stem response to speech: a biological marker of auditory processing. Ear Hearing 26:424–434

Johnson KL, Nicol T, Zecker SG, Bradlow AR, Skoe E, Kraus N (2008) Brainstem encoding of voiced consonant-vowel stop syllables. Clin Neurophysiol 119(11):2623–2635

Johnstone RA (1995) Sexual selection, honest advertisement and the handicap principle – reviewing the evidence. Biol Rev 70(1):1–65

Joris PX, Schreiner CE, Rees A (2004) Neural processing of amplitude-modulated sounds. Physiol Rev 84:541–577. doi:10.1152/physrev.00029.2003

Kidd G Jr, Mason CR, Richards VM, Gallun FJ, Durlach NI (2007) Informational masking. In: Yost WA, Popper AN, Fay RR (eds) Auditory perception of sound sources. Springer, New York, pp. 143–190

Konishi M (1970) Comparative neurophysiological studies of hearing and vocalizations in songbirds. Z Vgl Physiol 66:257–272

Köppl C, Gleich O, Manley GA (1993) An auditory fovea in the barn owl. J Comp Physiol A 171:695–704

Krishnan A, Gandour JT (2009) The role of the auditory brainstem in processing linguistically-relevant pitch patterns. Brain Lang 110(3):135–148

Kroodsma DE, Miller EH (1996) Ecology and evolution of acoustic communication in birds. Cornell University Press, Ithaca

Kuwada S, Batra R, Maher VL (1986) Scalp potentials of normal and hearing-impaired subjects in response to sinusoidally amplitude-modulated tones. Hear Res 21:179–192. doi:10.1016/0378-5955(86)90038-9

Lehongre K, Del Negro C (2011) Representation of the bird's own song in the canary HVC: contribution of broadly tuned neurons. Neuroscience 173:93–109

Lohr B, Dooling RJ (1998) Detection of changes in timbre and harmonicity in complex sounds by zebra finches (*Taeniopygia guttata*) and budgerigars (*Melopsittacus undulatus*). J Comp Psychol 112:36–47

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(4):763–777

Lucas JR, Freeberg TM, Krishnan A, Long GR (2002) A comparative study of avian auditory brainstem responses: correlations with phylogeny and vocal complexity, and seasonal effects. J Comp Physiol A 188:981–992

Lucas JR, Freeberg TM, Long GR, Krishnan A (2007) Seasonal variation in avian auditory evoked responses to tones: a comparative analysis of Carolina chickadees, tufted titmice, and white-breasted nuthatches. J Comp Physiol A 193:201–215

Lucas JR, Vélez A, Henry KS (2015) Habitat-related differences in auditory processing of complex tones and vocal signal properties in four songbirds. J Comp Physiol A 201:395–410. doi:10.1007/s00359-015-0986-7

Lynch KS, Wilczynski W (2005) Gonadal steroids vary with reproductive stage in a tropically breeding female anuran. Gen Comp Endocr 143:51–56

Lynch KS, Wilczynski W (2008) Reproductive hormones modify reception of species-typical communication signals in a female anuran. Brain Behav Evol 71:143–150

Lynch KS, Stanely Rand A, Ryan MJ, Wilczynski W (2005) Plasticity in female mate choice associated with changing reproductive states. Anim Behav 69:689–699

Lynch KS, Crews D, Ryan MJ, Wilczynski W (2006) Hormonal state influences aspects of female mate choice in the túngara frog (*Physalaemus pustulosus*). Horm Behav 49:450–457

Maddison CJ, Anderson RC, Prior NH, Taves MD, Soma KK (2012) Soft song during aggressive interactions: seasonal changes and endocrine correlates in song sparrows. Horm Behav 62 (4):455–463

Marean GC, Burt JM, Beecher MD, Rubel EW (1998) Auditory perception following hair cell regeneration in European starling (*Sturnus vulgaris*): Frequency and temporal resolution. J Acoust Soc Am 103:3567–3580

Marler P, Slabbekoorn H (eds) (2004) Nature's Music: the science of bird song. Elsevier Academic, San Diego

Marten K, Marler P (1977) Sound transmission and its significance for animal vocalizations: I. Temperate habitats. Behav Ecol Sociobiol 2:291–302

Miller CT, Bee MA (2012) Receiver psychology turns 20: is it time for a broader approach? Anim Behav 83:331–343

Mills JH, Schmiedt RA, Schulte BA, Dubno JR (2006) Age-related hearing loss: a loss of voltage, not hair cells. Semin Hear 27:228–236

Moore BCJ (1993) Frequency analysis and pitch perception. In: Yost W, Popper A, Fay R (eds) Human psychophysics. Springer, New York, pp. 58–89

Morton ES (1975) Ecological sources of selection on avian sounds. Am Nat 109:17–34

Mostrom AM, Curry RL, Lohr B (2002) Carolina Chickadee (*Poecile carolinensis*). In: Poole A (ed) The birds of North America online. Cornell Lab of Ornithology, Ithaca

Musacchia G, Sams M, Skoe E, Kraus N (2007) Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. Proc Natl Acad Sci USA 104(40):15894–15898

Niven JE, Laughlin SB (2008) Energy limitation as a selective pressure on the evolution of sensory systems. J Exp Biol 211:1792–1804

Njegovan M, Weisman R (1997) Pitch discrimination in field- and isolation-reared black-capped chickadees (*Parus atricapillus*). J Comp Psychol 111(3):294–301

Noirot IC, Adler HJ, Cornil CA, Harada N, Dooling RJ, Balthazart J, Ball GF (2009) Presence of aromatase and estrogen receptor alpha in the inner ear of zebra finches. Hear Res 252:49–55

Ohashi T, Ochi K, Nishino H, Kenmochi M, Yoshida K (2005) Recovery of human compound action potential using a paired-click stimulation paradigm. Hear Res 203:192–200. doi:10.1016/j.heares.2004.12.001

Parham K, Zhao HB, Kim DO (1996) Responses of auditory nerve fibers of the unanesthetized decerebrate cat to click pairs as simulated echoes. J Neurophysiol 76:17–29

Parthasarathy A, Bartlett EL (2011) Age-related auditory deficits in temporal processing in F-344 rats. Neuroscience 192:619–630

Patterson RD (1976) Auditory filter shapes derived with noise stimuli. J Acoust Soc Am 59:640–654. doi:10.1121/1.380914

Phillmore LS, Sturdy CB, Weisman RG (2003) Does reduced social contact affect discrimination of distance cues and individual vocalizations? Anim Behav 65:911–922

Pichora-Fuller MK, Souza PE (2003) Effects of aging on auditory processing of speech. Int J Audiol 42:S11–S16

Popov VV, Supin AY, Wang D, Wang KX (2006) Nonconstant quality of auditory filters in the porpoises, *Phocoena phocoena* and *Neophocaena phocaenoides* (Cetacea, Phocoenidae). J Acoust Soc Am 119:3173–3180

Richards DG, Wiley RH (1980) Reverberations and amplitude fluctuations in the propagation of sound in a forest: Implications for animal communication. Am Nat 115:381–399

Ríos-Chelén AA, Lee GC, Patriccelli GL (2015) Anthropogenic noise is associated with changes in acoustic but not visual signals in red-winged blackbirds. Behav Ecol Sociobiol 69:1139–1151

Ritchison G (1983) Vocalizations of the white-breasted nuthatch. Wilson Bull 95:440–451

Rohmann KN, Bass AH (2011) Seasonal plasticity of auditory hair cell frequency sensitivity correlates with plasma steroid levels in vocal fish. J Exp Biol 214:1931–1942

Rohmann KN, Deitcher DL, Bass AH (2009) Calcium-activated potassium (BK) channels are encoded by duplicate slo1 genes in teleost fishes. Mol Biol Evol 26:1509–1521

Rohmann KN, Fergus DJ, Bass AH (2013) Plasticity in ion channel expression underlies variation in hearing during reproductive cycles. Curr Biol 23:678–683

Rohmann KN, Tripp JA, Genova RM, Bass AH (2014) Manipulation of BK channel expression is sufficient to alter auditory hair cell thresholds in larval zebrafish. J Exp Biol 217:2531–2539

Ronald KL, Fernandez-Juricic E, Lucas JR (2012) Taking the sensory approach: how individual differences in sensory perception can influence mate choice. Anim Behav 84(6):1283–1294

Russo NM, Nicol TG, Zecker SG, Hayes EA, Kraus N (2005) Auditory training improves neural timing in the human brainstem. Behav Brain Res 156:95–103

Sanes DH, Woolley SMN (2011) A behavioral framework to guide research on central auditory development and plasticity. Neuron 72(6):912–929

Saunders JC, Coles RB, Gates GR (1973) The development of auditory evoked responses in the cochlea and cochlear nuclei of the chick. Brain Res 63:59–74

Schrode KM, Bee MA (2015) Evolutionary adaptations for the temporal processing of natural sounds by the anuran peripheral auditory system. J Exp Biol 218:837–848

Searcy WA, Nowicki A (2005) The evolution of animal communication. Princeton University Press, Princeton

Sisneros JA (2009) Adaptive hearing in the vocal plainfin midshipman fish: getting in tune for the breeding season and implications for acoustic communication. Integr Zool 4:33–42

Sisneros JA, Forlano PM, Deitcher DL, Bass AH (2004) Steroid-dependent auditory plasticity leads to adaptive coupling of sender and receiver. Science 305:404–407

Slabbekoorn H, Peet M (2003) Birds sing at higher pitch in urban noise. Nature 424:267

Slabbekoorn H, Smith TB (2002) Bird song, ecology, and speciation. Philos Trans R Soc Lond B Biol Sci 357:493–503

Sockman KW, Gentner TQ, Ball GF (2002) Recent experience modulates forebrain gene-expression in response to mate-choice cues in European starlings. Proc Biol Sci 269:2479–2485

Sockman KW, Gentner TQ, Ball GF (2005) Complementary neural systems for the experience-dependent integration of mate-choice cues in European starlings. J Neurobiol 62:72–81

Suga N, Neuweiler G, Möller J (1976) Peripheral auditory tuning for fine frequency analysis by CF-FM bat, *Rhinolophus ferrumequinum*. 4. Properties of peripheral auditory neurons. J Comp Physiol 106:111–125

Supin AY, Popov VV (1995) Temporal resolution in the dolphin's auditory system revealed by double-click evoked potential study. J Acoust Soc Am 97:2586–2593. doi:10.1121/1.411913

Swaminathan J, Mason CR, Streeter TM, Best V, Kidd G Jr, Patel AD (2015) Musical training, individual differences and the cocktail party problem. Sci Rep 5:14401

Vasconcelos RO, Sisneros JA, Amorim MCP, Fonseca PJ (2011) Auditory saccular sensitivity of the vocal Lusitanian toadfish: low frequency tuning allows acoustic communication throughout the year. J Comp Physiol A 197:903–913

Vélez A, Gall MD, Fu J, Lucas JR (2015a) Song structure, not high-frequency song content, determines high-frequency auditory sensitivity in nine species of New World sparrows (Passeriformes: Emberizidae). Funct Ecol 29:487–449

Vélez A, Gall MD, Lucas JR (2015b) Seasonal plasticity in auditory processing of the envelope and temporal fine structure of sounds in three songbirds. Anim Behav 103:53–63. doi:10.1016/j.anbehav.2015.01.036

Wiley RH (1991) Associations of song properties with habitats for territorial oscine birds of eastern North America. Am Nat 138(4):973–993

Wiley RH, Richards DG (1978) Physical constraints on acoustic communication in atmosphere – implications for evolution of animal vocalizations. Behav Ecol Sociobiol 3:69–94

Wong A, Gall MD (2015) Frequency sensitivity in the auditory periphery of male and female black-capped chickadees (*Poecile atricapillus*). Zoology 118(5):357–363

Wong RY, So P, Cummings ME (2011) How female size and male displays influence mate preference in a swordtail. Anim Behav 82:691–697

Woolley SMN (2012) Early experience shapes vocal neural coding and perception in songbirds. Dev Psychobiol 54(6):612–631

Woolley SMN, Gill PR, Fremouw T, Theunissen FE (2009) Functional groups in the avian auditory system. J Neurosci 29:2780–2793

Zahavi A (1975) Mate selection – a selection for a handicap. J Theor Biol 53:205–214

Zeng S-J, Székely T, Zhang X-W et al (2007) Comparative analyses of song complexity and song-control nuclei in fourteen oscine species. Zoolog Sci 24:1–9. doi:10.2108/zsj.24.1

Chapter 3
Perceptual and Neural Mechanisms of Auditory Scene Analysis in the European Starling

Georg M. Klump

Abstract Humans and many other animals, such as songbirds, communicate acoustically in large, dense social groups. In such environments, the signals produced by different signalers commonly overlap in time and frequency, and background noise can be intense. How can receivers make sense of the acoustic scene when there is so much noise and acoustic clutter? The answer is that vocal communication in such environments engages a suite of perceptual and cognitive mechanisms responsible for parsing the acoustic scene into perceptually discrete auditory "objects" or "streams" of behavioral relevance. In this chapter, I review psychophysical and neurophysiological studies of European starlings (*Sturnus vulgaris*, Sturnidae) that have aimed to identify mechanisms underlying the perceptual organization of complex acoustic scenes. The focus of this review is on recent efforts to discover neural mechanisms for auditory scene analysis (ASA) that promote signal detection (e.g., comodulation masking release and the comodulation detection difference), signal recognition (e.g., perceptual restoration), and signal segregation (e.g., auditory streaming) under adverse listening conditions. The chapter emphasizes that key insights into the neural codes for ASA are to be gained by integrating neurophysiological approaches with objective measures of psychophysical performance in animal models for which receiving communication signals in a crowd is a key feature of their biology.

G.M. Klump (✉)
Cluster of Excellence Hearing4all, Animal Physiology and Behaviour Group, Department for Neuroscience, School of Medicine and Health Sciences, University of Oldenburg, Oldenburg, Germany
e-mail: georg.klump@uol.de

M.A. Bee, C.T. Miller (eds.), *Psychological Mechanisms in Animal Communication*, Animal Signals and Communication 5,
DOI 10.1007/978-3-319-48690-1_3

3.1 Introduction

A dawn chorus of birds, or the cacophony of a night roost, with thousands of birds singing simultaneously, constitute challenging acoustic environments that rival human cocktail parties (Bee and Micheyl 2008; Chap. 4). For a songbird like the European starling (*Sturnus vulgaris*), critical social and communicative interactions depend on accurately receiving vocal signals in such environments. It has been suggested, for example, that starlings at their night roost share song elements with their neighbors (Hausberger et al. 2008). Furthermore, studies on responses of starling females to males' songs suggest that females choose their mates in the breeding colony based on listening to features of their songs (Eens et al. 1991; Mountjoy and Lemon 1996; Gentner and Hulse 2000). Starling song is composed of long sequences of different element types, with the number of elements (also called motifs) increasing with age and reflecting the bird's condition (e.g., Eens 1992a, b). These different motifs are composed of complex and stereotyped spectro-temporal patterns that are presented in a hierarchical syntactic sequence (Eens et al. 1989; Adret-Hausberger and Jenkins 1988; Comins and Gentner 2014). Starlings also have whistle-like signals that are well suited for interaction and individual recognition over long distances and, thus, must be especially adapted to be easily detected and have features that are robust against degradation in sound transmission (e.g., Hausberger et al. 2008). These examples demonstrate that starling receivers have to solve complex problems of *auditory scene analysis (ASA)* (Bregman 1990) that, in many respects, resemble the problems humans have to solve when we listen to speech in noisy social settings. In addition, starlings, being a typical songbird, are vocal learners that, similar to humans, acquire their vocalizations by learning from conspecifics in their social environment (e.g., Chaiken et al. 1993, 1994). Song learning changes the neural representation of song motifs in the starling forebrain (e.g., Thompson and Gentner 2010; Jeanne et al. 2011; George and Cousillas 2013). Since there are so many parallels between human and bird vocal communication (Soha and Peters 2015), we stand to achieve a much better understanding of ASA in humans, and perhaps other animals too, by studying its mechanisms in songbirds like the European starling (see also Hulse 2002).

3.2 The Neuronal Substrate of Auditory Processing in Birds

Auditory perceptual mechanisms rely on hierarchical neuronal processing of stimulus features in the auditory pathway. This pathway, originating in the inner ear and reaching up to the cortex, is often referred to as the bottom-up pathway (for a schematic of the bird auditory pathway, see Carr and Code 2000). On the basilar papilla of the starling inner ear, the frequency components of sound are represented tonotopically, with low frequencies exciting the apical region and high frequencies exciting the base (e.g., Gleich and Manley 1988; Gleich 1989). This tonotopy is preserved throughout the auditory pathway up to the starling's cortex. The spiral ganglion neurons that innervate the hair cells of the sensory epithelium of the inner ear and form the auditory nerve not only represent frequency by the place they innervate along the basilar papilla (i.e., forming a labeled-line code of frequency), but they can also represent frequency by the time interval between action potentials that are phase locked to the period of the stimulus (i.e., provide a temporal code of frequency). Phase locking, however, is limited to frequencies below about 2 kHz (Gleich and Narins 1988). Above that limit, action potentials can still be locked to the temporal structure of the sound. However, the neurons then fire correlated with the envelope of the sound rather than with the carrier frequency (Gleich and Klump 1995). Henry et al. (Chap. 2) discuss how signal processing by these neural mechanisms can vary across seasons and among species and individuals in songbirds.

Avian auditory nerve fibers project to the cochlear nucleus complex, with the nucleus angularis representing mainly the frequency spectrum of sound and the nucleus magnocellularis mainly representing the temporal structure of sound (Carr and Code 2000). The first binaural comparison of the temporal pattern in the neuronal activity originating from the left and the right ears occurs in the nucleus laminaris, which computes interaural time differences, a cue that provides information about the direction of a sound source in the horizontal plain. Binaural comparison of interaural intensity differences, the second major cue for sound localization in the horizontal plain, is provided higher up in the auditory pathway, presumably by the nuclei of the lateral lemniscus. Note that in the starling, interaural time differences may be enhanced by a physical binaural interaction of the sounds from both ears at the tympana, which provides the bird with a directional pressure-difference system (Klump and Larsen 1992). The bird's inferior colliculus, in the midbrain, receives input from both the left and right auditory pathways. It not only represents cues related to the direction of sound sources but also other cues related to the spectro-temporal structure of the sound. Inferior colliculus neurons in a close relative of the starling, the myna bird (*Gracula religiosa intermedia*), are tuned to the modulation frequencies in the envelope of sound, thus providing for another cue representing sound sources (Hose et al. 1987). In the auditory thalamus, Bigalke-Kunz et al. (1987) observed neurons that were tonotopically organized and that not only had a narrowband frequency tuning (partly created by inhibitory sidebands bordering the excitation), but that were also able to faithfully represent the temporal patterns of modulated sounds by their temporally structured ongoing response. Similar responses were observed in the thalamic projection area in the starling auditory forebrain, which also has a clear tonotopic organization (Rübsamen et al. 1986) and an ability to represent the spectro-temporal patterns in the stimulus by a temporally patterned response (Knipschild et al. 1992). Starling secondary auditory forebrain areas represent more complex features of sounds and provide a neuronal substrate for recognition of song elements and for acquiring templates relevant in song learning (e.g., Gentner and Margoliash 2003; Thompson and Gentner 2010; Jeanne et al. 2011).

While bottom-up processing in the starling brain has been studied with respect to many stimulus features, top-down, cognitive effects on processing in the auditory pathway have received less attention. The main reason for this lack of studies is the fact that top-down effects reflect processing in active listening subjects, such as a bird engaged in a certain natural behavior or in a psychoacoustic listening task. Studying these effects at a neurophysiological level requires recording from behaving subjects in conjunction with recording from one or more sites in the brain, which is difficult to achieve. In the starling, modification of processing of sounds in relation to the animal's behavioral state has been observed with respect to the processing of learned song elements (e.g., Gentner and Margoliash 2003; Knudsen and Gentner 2013), indicating a role for top-down processing in hearing and sound communication.

3.3 Principles of Auditory Scene Analysis

In his seminal book on ASA, Bregman (1990) differentiated between "primitive" mechanisms and "schema-based" mechanisms, the former representing bottom-up processes and the latter top-down processes in ASA. These mechanisms provide for a separated neural representation of sounds from different sources in the brain while binding the representations of sounds from the same source together (see also Chap. 4). Sounds originating from a specific source are referred to as "objects" or, if presented in a sequence, "auditory streams." Primitive mechanisms for ASA rely on the physical laws of acoustics. Sound-generating structures will automatically provide relevant cues for primitive ASA mechanisms. For example, if a sound source generates a rapid series of brief sound pulses (such as the human larynx producing vowels in speech), a harmonic complex is produced in which the fundamental frequency is the inverse of the pulse rate. In such sounds, the harmonic relatedness of the components, or the sound's "harmonicity," can serve as a cue to the fact that the components originate from a common source. Furthermore, vocal mechanisms that gate a sound (such as the human tongue or lips producing plosives in speech) create a common onset for different frequency components that serves as a cue indicating their common source. Vocal mechanisms will rarely produce sounds with large, instantaneous changes in frequency (human yodeling being an exception), which makes the similarity in frequency of successive sounds a reliable cue that sequential sounds originate from one source. Additional cues that are useful for characterizing sound sources include more complex spectral cues (e.g., spectral shape resulting in a specific timbre, e.g. see Bregman et al. 2016) or cues related to the spatial location of a sound source. Finally, the transmission of sounds from the source to the listener also creates cues that are characteristic for each source (see Sect. 3.4).

Schema-based ASA mechanisms more commonly rely on higher-level, cognitive processes, such as auditory memory, attention, and learned expectations, although in the case of short-term memory, primitive mechanisms such as stimulus-specific adaptation may still prevail (Khouri and Nelken 2015).

For example, short-term memory provides the basis for the operation of an “old-plus-new” heuristic in ASA (Bregman 1990). The operation of such a heuristic is evidenced when a deviation from the regularities in the previous sequence or prior stimulus elicits a percept of a new object. Related effects have been observed in the oddball paradigm used in the study of the mismatch negativity in the electroencephalogram (EEG) signal (e.g., Näätänen et al. 2005). Thus, some apparent short-term, memory-related effects may still result from bottom-up mechanisms that rely on neuronal adaptation. Long-term auditory memory effects, however, can more convincingly be associated with top-down, schema-based mechanisms. An example for such effects of auditory templates stored in memory is given for the perceptual restoration of speech sounds, which is affected by the knowledge of the speech (e.g., Samuel 1996; similar effects are observed in starlings, see Sect. 3.5). The observation that naïve young songbirds that have never heard a conspecific sing nevertheless show a predisposition for learning the song of their own species over that from other species indicates that such templates can even have an evolutionary basis (reviewed in Soha and Peters 2015; see also Chap. 4). Finally, it has been demonstrated in human subjects that instructions in an ASA paradigm can affect perception by focusing a listener’s attention on specific features of the sound (e.g., van Noorden 1975). It may well be that some loud and prominent introductory notes, such as the whistles at the start of a starling song bout, serve a similar purpose.

3.4 Auditory Scene Analysis and Signal Detection

Different sound sources in the natural environment are rarely synchronized. Typically, animals wait for a quiet interval in the background noise produced, for example, by conspecifics before they vocalize themselves (Ficken et al. 1974; Zelick and Narins 1982; Chap. 10). Bird songs lasting a few seconds have a distinct temporal structure, and counter-singing birds only partially overlap their competitor’s song, if at all (reviewed in Todt and Naguib 2000). As a result, the bird dawn chorus that is typical for the breeding season in temperate habitats, and that also produces a masking background for communication, has distinct level fluctuations (Nelken et al. 1999; see Vélez and Bee 2011 for parallel analyses of frog choruses). These fluctuations in the level of background noise are temporally incoherent with respect to the temporal structure of the song that a particular singer broadcasts in such a situation. That is, signals from one source and the background noise produced by other sources have distinct and independent spectro-temporal patterns. In addition to the level fluctuations produced by the sources themselves, the turbulences on the path of transmission of a signal from the sound source to the listener will modulate the signal envelope in a way that is specific for the route taken by the signal. Signals from other sources having a different path of transmission will have different patterns of modulations imposed on them. As a result of the sum of these processes, signals and background noise have temporally independent (i.e., uncorrelated) patterns of level variation that can be exploited for improving signal detection. Two experimental paradigms have been used in laboratory studies to investigate these effects on the detectability of signals: the *comodulation masking release (CMR)* paradigm and the *comodulation detection difference (CDD)* paradigm.

3.4.1 Comodulation Masking Release

In a CMR paradigm (Figs. 3.1 and 3.2), an unmodulated signal with a constant level, such as a pure tone, has to be detected in a modulated noise background with a fluctuating level and a spectrum that, at least partially, overlaps that of the signal. The release from masking attributable to CMR corresponds to the difference in signal detection threshold between a condition in which different frequency bands of the masker have uncorrelated level variations and a condition in which the level variations in different frequency bands of the masker are correlated. Sounds in different frequency ranges that have correlated envelope modulations are said to be "comodulated." In comodulated conditions, the background noise is characterized by distinctive peaks and dips in the envelope since level fluctuations across different frequency bands occur simultaneously (see waveform in Fig. 3.1b). In uncorrelated conditions, the masker envelope lacks high peaks and deep dips because the amplitude is rarely simultaneously high or low across all frequency bands of the masker (Fig. 3.1a). Thus, the background noise appears to be unmodulated, or at least less modulated, than in the comodulated condition (cf. Fig. 3.1a, b). The term CMR itself is derived from the observation that thresholds for signal detection in comodulated conditions are typically lower than in unmodulated or uncorrelated conditions. That is, there is a so-called release from masking in comodulated conditions compared with unmodulated or uncorrelated conditions.

Two variants of the CMR paradigm have been investigated. In the first paradigm, a noise masker with varying bandwidth is centered on the tone, and an increasing CMR with increasing masker bandwidth is observed (Fig. 3.1c, d). This paradigm is referred to as the "band-widening paradigm" in the literature (e.g., Hall et al. 1984; Schooneveldt and Moore 1989). In the second paradigm, referred to as the "flanking-band paradigm" (e.g., Schooneveldt and Moore 1987), one narrowband noise is centered on the signal frequency, and additional flanking bands composed of narrowband noise are presented at remote frequencies (Fig. 3.2a, b). Similar to the band-widening paradigm, the masker envelope has greater peakiness in the correlated condition than in the alternate (uncorrelated) condition (cf. waveforms in Fig. 3.2a, b).

CMR has been investigated in behavioral experiments with starlings, and the related effects in the responses of starling forebrain neurons have also been described. When studying CMR in starlings with the band-widening paradigm, Klump and Langemann (1995) observed a perceptual masking release that was

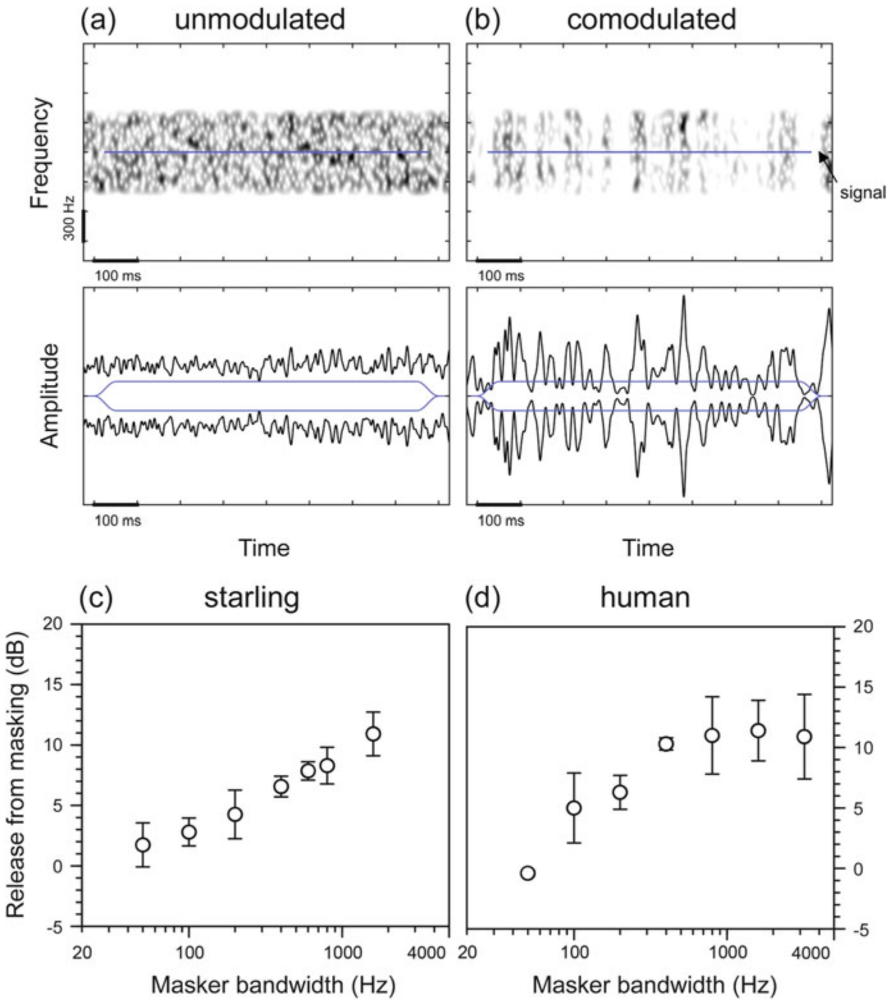


Fig. 3.1 Comodulation masking release (CMR) in starlings and humans measured in a band-widening paradigm. Panels (**a**) and (**b**) depict spectrograms (*top*) and waveforms (*bottom*) of stimuli used in the unmodulated and comodulated conditions, respectively. In the waveform depictions, the temporal envelopes of signals and maskers are outlined in *blue* and *black*, respectively. In this paradigm, CMR corresponds to the improvement in threshold for detecting an unmodulated signal, such as a pure tone, in the presence of (**b**) a comodulated masker compared with (**a**) an unmodulated masker having the same bandwidth. The masker in (**b**) is created by multiplying the unmodulated noise band shown in (**a**) with a low-pass noise [e.g., with a cutoff frequency of 50 Hz, as used to collect the data in panels (**c**) and (**d**)]. CMR is, thus, the difference between the threshold for signal detection in the unmodulated masker and the comodulated masker, where positive values indicate CMR. CMR increases with increasing bandwidth and decreasing rate of envelope fluctuation. Data in (**c**) and (**d**) show the mean (±s.d.) magnitude of the CMR effect in starlings (redrawn from Klump and Langemann 1995) and humans (redrawn from Schooneveldt and Moore 1989), respectively, as a function of masker bandwidth

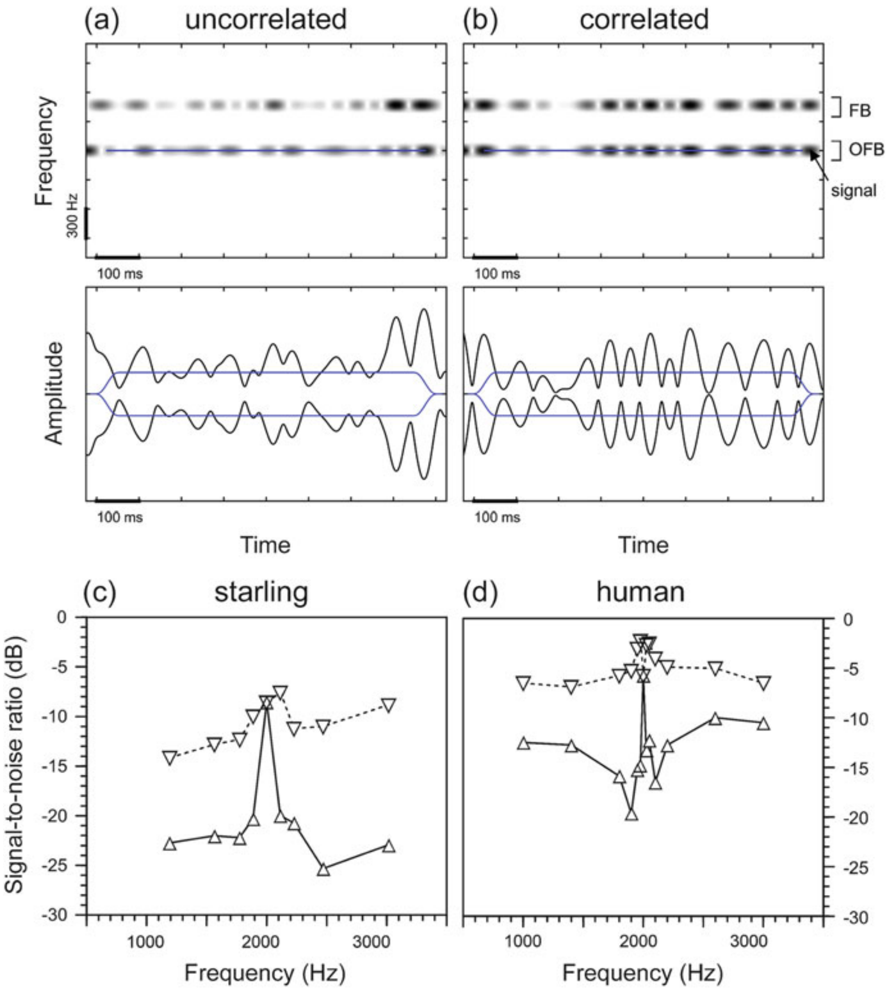


Fig. 3.2 Comodulation masking release (CMR) in starlings and humans measured in a flanking-band paradigm. Panels (**a**) and (**b**) depict spectrograms (*top*) and waveforms (*bottom*) of stimuli used in the uncorrelated and correlated conditions, respectively. In the waveform depictions, the temporal envelopes of signals and maskers (summed across noise bands) are outlined in *blue* and *black*, respectively. In this paradigm, one narrow band of masking noise is centered on the signal frequency (on-frequency band, OFB), and a second narrowband noise is presented at a frequency remote from that of the signal (flanking band, FB). Detection threshold for a signal presented in OFB and FB maskers with (**a**) uncorrelated envelopes is compared to the detection threshold of the same signal in OFB and FB maskers with (**b**) correlated (i.e., similar) envelopes. Detection thresholds are lower in maskers with correlated envelopes [*white up-pointing triangle* in panels (**c**) and (**d**)] than in maskers with uncorrelated envelopes [*white down-pointing triangle* in panels (**c**) and (**d**)]. The magnitude of CMR corresponds to the threshold difference between these two conditions. Data in (**c**) and (**d**) show the mean detection thresholds in starlings (redrawn from Klump et al. 2001) and humans (redrawn from Schooneveldt and Moore 1987), respectively, as a function of the center frequency of the FB

similar to that observed in humans obtained with similar stimuli (cf. Fig. 3.1c, d). Subject birds were trained using a go/no-go paradigm to detect a 2-kHz tone that was spectrally centered in a band-limited masking noise. The masking noise was either unmodulated (Fig. 3.1a) or comodulated (Fig. 3.1b). Across different conditions, the bandwidth of the masker varied between 50 and 1600 Hz. As illustrated in Fig. 3.1c, the magnitude of CMR observed, measured as the improvement in masked signal detection thresholds in the comodulated conditions relative to the unmodulated conditions, increased from 1.6 to 11.8 dB as the masker bandwidth increased from 50 to 1600 Hz. Stated another way, the starlings could detect the target tone at thresholds that were up to about 12 dB lower (i.e., better) in the comodulated condition relative to the unmodulated condition. These results for starlings were strikingly similar to earlier results from studies of human listeners reported by Schooneveldt and Moore (1989; cf. Fig. 3.1c, d). In a second experiment, Klump and Langemann (1995) showed that the magnitude of CMR was inversely correlated with the rate and peakiness of the temporal fluctuations in the masker envelope, which they manipulated by varying the bandwidth of the modulator used to impose envelope fluctuations on the masker. The magnitude of CMR was highest when the comodulated masker fluctuated at slow rates ($<$12.5 Hz).

A similar correspondence in perception between starlings and humans was also observed in studies of CMR using the flanking-band paradigm (Langemann and Klump 2001; Klump et al. 2001). This is illustrated in Fig. 3.2c, showing data from a go/no-go task in which starlings were trained to detect a 2-kHz tone in the presence of an *on-frequency band (OFB)* masker centered at 2 kHz (Klump et al. 2001). When a *flanking band (FB)* at the frequency indicated along the x-axis was added to the OFB located at 2 kHz, thresholds were significantly lower when the FB and OFB had envelopes that were correlated (△ symbols in Fig. 3.2c) compared to when the two noise bands fluctuated independently and, thus, had uncorrelated envelopes (▽symbols in Fig. 3.2c). The magnitude of CMR, computed as the difference in thresholds between the correlated and uncorrelated conditions, was typically on the order of 10 dB or more, depending on the location of the FB relative to the OFB (Fig. 3.2c). These measurements of CMR were, again, strikingly similar to data from human listeners (cf. Fig. 3.2c, d; Schooneveldt and Moore 1987). Langemann and Klump (2001) trained starlings in another go/no-go task to detect a short, 2-kHz tone pip presented in a masker composed of between one and five narrowband noises, one of which was centered on 2 kHz. The envelope of each noise band fluctuated with sinusoidal modulation at a rate of 10 Hz. In correlated conditions, the envelopes of all noise bands fluctuated in phase to create comodulation; in uncorrelated conditions, the envelopes of adjacent noise bands were phase shifted relative to each other by +90° to create incoherence. Compared with uncorrelated conditions, masked thresholds in the correlated conditions improved by about 26 dB for maskers with more than one flanking band, again demonstrating a substantial CMR in starlings. Together, these behavioral studies of CMR in starlings indicate that the birds are adept at exploiting temporal correlations in background noise to improve signal detection.

Electrophysiological data based on single-unit and multiunit recordings made from awake starlings using radiotelemetry have identified neuronal correlates of CMR in the auditory forebrain using both the band-widening paradigm (Klump and Nieder 2001) and the flanking-band paradigm (Nieder and Klump 2001; Hofer and Klump 2003). For example, Hofer and Klump (2003) used the flanking-band paradigm to measure neural detection thresholds in response to a pure tone presented at a neuron's best excitatory frequency. The OFB was also centered at the neuron's best excitatory frequency, and the FB was presented at a remote frequency that was either in an excitatory or inhibitory portion of the neuron's frequency tuning curve. Across conditions, a neural CMR of about 3–7 dB in magnitude was observed when the OFB and FB were correlated, compared to an uncorrelated condition. This is somewhat smaller than the magnitude of CMR observed in psychophysical experiments. However, across studies, an average of about 20 % of the recording sites, and in some stimulus conditions up to about 60 %, showed a release from masking in comodulated or correlated conditions that was similar to, or even exceeded in magnitude, the corresponding behavioral measures of CMR. Thus, the responses of starling forebrain neurons could account well for the starling's perceptual CMR.

3.4.2 The Comodulation Detection Difference

The CDD paradigm (Fig. 3.3) differs from the CMR paradigm in that the signal envelope is also modulated (e.g., McFadden 1987). In a typical CDD experiment (as was the case in some CMR experiments), the masker is composed of a number of narrowband noises that fluctuate in level (Fig. 3.3). These fluctuations can occur independently across masker noise bands (Fig. 3.3a), or they can be correlated across masker noise bands (Fig. 3.3b, c). In addition, the fluctuations in the masker noise bands can have one of three relationships with those in the signal. They can all fluctuate independently of the fluctuations in the signal (Fig. 3.3a, the all-uncorrelated condition), they can be correlated with each other and with the fluctuations in the signal (Fig. 3.3b, the all-correlated condition), or they can be correlated with each other but fluctuate independently of the fluctuations in the signal (Fig. 3.3c, the co-uncorrelated condition).

While the summed envelope of all masker bands has distinctive peaks and dips in the all-correlated and co-uncorrelated conditions, the summed envelope of the masker in the all-uncorrelated condition is less peaky. Thus, in comparing CMR and CDD experiments, the uncorrelated (CMR) and all-uncorrelated (CDD) masker conditions correspond to each other, and the comodulated (CMR) and the co-uncorrelated (CDD) masker conditions correspond to each other. A key difference between the two experiments is that the signal has a steady envelope in CMR (Figs. 3.1 and 3.2) and a modulated envelope in CDD (Fig. 3.3). The all-correlated condition in the CDD paradigm is special in that both the signal and masker noise bands all have highly correlated envelopes. It has been suggested that both temporal

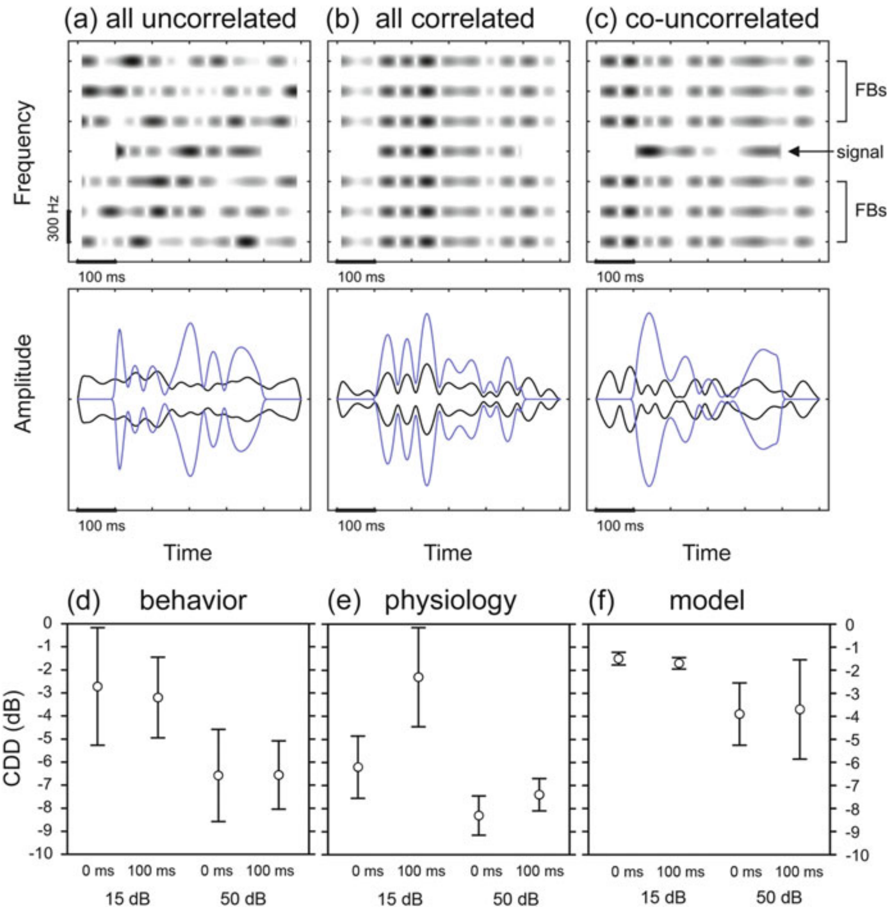


Fig. 3.3 The comodulation detection difference (CDD) and its neural correlates in starlings. Panels (**a**–**c**) depict spectrograms (*top*) and waveforms (*bottom*) of stimuli used in the all-uncorrelated, all-correlated, and co-uncorrelated conditions, respectively. In the waveform depictions, the temporal envelopes of signals and maskers (summed across noise bands) are outlined in *blue* and *black*, respectively. The CDD represents the fact that detection of a modulated narrowband noise signal is easier in a masker composed of flanking bands (FB) with correlated envelopes that differ from the modulation pattern of the signal envelope [the co-uncorrelated condition in panel (**c**)] than in a masker composed of flanking bands with envelopes that are correlated with each other and with the signal envelope [the all-correlated condition in panel (**b**)]. Panels (**d**–**f**) show the size of the CDD effect in response to the stimuli shown in (**b**) and (**c**) that was observed (**d**) in starling perception, (**e**) in the rate response of starling auditory forebrain neurons, and (**f**) as predicted by a model of the auditory periphery. Although the model was able to represent the details of the rate-intensity functions for a signal presented with increasing level in a constant level masker, it could not predict the full magnitude of the CDD effect, suggesting that more central processes play a role (for details, see Langemann and Klump 2007; Bee et al. 2007)

masking and auditory grouping effects contribute to the perception of signals in maskers that are modulated across different frequency bands (Borrill and Moore 2002; Verhey et al. 2003).

Similar to studies of CMR in starlings, a close correspondence between perception (Langemann and Klump 2007) and neural representation (Bee et al. 2007) has been found in starlings using the CDD paradigm. In behavioral experiments, starlings were trained in a go/no-go task to detect a 400-ms narrowband noise signal in the presence of flanking bands composed of six 600-ms narrowband noise maskers (Langemann and Klump 2007). The envelopes of signals and flanking bands were manipulated to create all-uncorrelated (Fig. 3.3a), all-correlated (Fig. 3.3b), and co-uncorrelated (Fig. 3.3c) conditions. Figure 3.3d depicts the magnitude of the CDD determined as the differences in threshold between the co-uncorrelated condition (Fig. 3.3c) and the all-correlated condition (Fig. 3.3b). Here, negative values indicate the relative improvement in thresholds in the co-uncorrelated condition, in which the envelopes of the flanking bands were correlated with each other, but uncorrelated with the envelope of the signal. Data are shown for separate replicates of the experiment in which the signal was delayed relative to the onset of the flanking bands by either 0 or 100 ms and in which the spectrum level of each flanking band was either 15 or 50 dB. Across conditions, thresholds were about 3–7 dB lower when the signal envelope fluctuated independently of the common envelope of the maskers compared to when all envelopes were correlated (Fig. 3.3d). Bee et al. (2007) investigated neural correlates of CDD in the auditory forebrains of awake starlings using the same stimuli as in the behavioral experiments. Neural signal detection thresholds were about 2–8 dB lower in the co-uncorrelated condition compared with the all-correlated condition (Fig. 3.3e). There was generally good correspondence between the magnitudes of behavioral and neural thresholds (cf. Fig. 3.3d, e). While neurophysiological recordings were limited to starling forebrain neurons, it is possible that at least part of the observed CDD effects in the forebrain (as well as the CMR effects discussed in Sect. 3.4.1) were due to processing at lower levels of the ascending auditory pathway. For example, a model that incorporated compressive mechanisms in the cochlea indicates that peripheral processing may substantially contribute to both CDD and CMR effects (Fig. 3.3f; Bee et al. 2007; Buschermöhle et al. 2007).

3.4.3 Relevance of CMR and CDD to Natural Listening Environments

How do the listening conditions created in laboratory studies of CMR and CDD reflect conditions in the natural environment? As pointed out above, background noise with a temporally fluctuating envelope is common in natural settings (e.g., Nelken et al. 1999; Vélez and Bee 2011). Bird songs, other animal vocalizations, and other natural sound ensembles also have a temporally distinct structure and a typical pattern of amplitude modulations (e.g., Singh and Theunissen 2003; Theunissen and Elie 2014). Sound propagation through an animal's natural habitat can also impose temporal envelope fluctuations on signals (see Fig. 3 in Wiley and Richards 1978). The envelope spectrum of these various natural sounds has most of its energy in the frequency range below 10 Hz, which is similar to the masking sounds used in the studies of CMR and CDD reviewed here. This suggests that processing by the bird's auditory system, as reflected by the CMR and CDD effects, may considerably improve detection and recognition of sounds in the animal's natural habitat. Thus, the laboratory studies reviewed in this section provide a representative account of the relevant effects for the perception of sounds in the natural environment.

3.5 Auditory Scene Analysis and Perceptual Restoration of Signals

In the visual domain, the so-called "picket-fence effect" is a well-known example of perceptual object completion. If we see an object (e.g., a dog) through a picket fence, we do not perceive the object (dog) as being sliced up and having missing pieces (i.e., the parts of the object for which the view is occluded by the fence). Rather, we perceive a complete object (dog) behind another complete object (picket fence). It is as if the brain "fills in" or perceptually restores the occluded parts such that we have no difficulty identifying a single, unified object behind the fence. A similar "filling-in" effect of perceptual restoration has been observed for the auditory system (reviewed in Warren 2008; Petkov and Sutter 2011). In humans, for example, when short tokens of speech are masked by a brief, loud sound, such as a cough, we often still perceive the speech as continuous through the cough, as if no part of the sentence had been masked. Studies have also shown that brief tokens of speech can even be completely removed and replaced with a burst of noise, yet listeners still report hearing the missing speech token. For speech sounds, such perceptual restoration is usually referred to as "phonemic restoration." Similar restorative effects occur for simpler, nonspeech sounds too. For example, we tend to hear a continuous frequency-modulated glide even when a brief central portion of the glide is removed and replaced with a short burst of masking noise. The ability to perceptually restore the missing parts of sound relies on bottom-up and possibly also top-down processing mechanisms.

Neurons in the afferent auditory pathway that are tuned to specific features of a sound signal could contribute to the bottom-up mechanism for restoration if they are sufficiently stimulated by the signal before and after the short time period in which only the masker is represented by the neuronal response. If the masker itself provides some stimulation to the neurons during the time period in which no signal is present, or if the neurons have a long response time constant, the brain can interpret the response of the neurons as the signal continuing through the masker. Such a mechanism has been demonstrated to operate both for pure tone signals in a wideband masker and for frequency-modulated signals (Sugita 1997; Petkov et al. 2007). In these two examples, neurons tuned to the frequency or the frequency modulation of the signal will respond as if the signal is ongoing, provided that some response during the masker alone time period is provided. Among mammals, behavioral studies of perceptual restoration in cats (*Felis catus*, Sugita 1997), cotton-top tamarins (*Saguinus oedipus*, Miller et al. 2001), rhesus macaques (*Macaca mulatta*, Petkov et al. 2003), and Mongolian gerbils (*Meriones unguiculatus*, Kobayasi et al. 2012) indicate that these nonhuman species also perceive restored signals. In contrast, studies of gray treefrogs (*Hyla chrysoscelis* and *Hyla versicolor*) and túngara frogs (*Physalaemus pustulosus*) have so far failed to find strong evidence of auditory perceptual restoration in anurans (Baugh et al. 2016; Seeba et al. 2010; Schwartz et al. 2010). What about birds?

Behavioral experiments with starlings indicate that these birds possess the ability for perceptual restoration of simple sounds like pure tones. In an operant conditioning task, starlings were trained to discriminate long tones from equivalent-duration stimuli composed of two short tones separated by a brief silent interval (Fig. 3.4a; Klump et al. 1999). The birds were required to peck one key when hearing the long tone and to peck a different key when hearing the two short tones. They were rewarded with food if they correctly reported the long tone. When they had mastered the discrimination, they were presented intermittently with stimuli in which the silent interval between the two short tones was filled with a broadband noise (Fig. 3.4a). At low levels of the noise (i.e., at high signal-to-noise ratios: S/N in Fig. 3.4b), the birds still reported hearing two short tones by pecking the appropriate key. At noise levels that were sufficiently high for masking the pure tone and making it inaudible (i.e., at lower S/N in Fig. 3.4b), the birds nevertheless pecked the key for indicating the presence of a long tone. They did not peck this key if presented with an isolated noise burst that was not bordered by two short tones. These results indicate that starlings are able to perceptually restore missing parts of simple acoustic stimuli. It was not necessary that the frequency band centered on the tone was present in the noise filling the silent gap between the two short tones. If the masker was loud enough, restoration occurred without spectral energy in the frequency band including the tone frequency. Such a loud masker results in a spread of masker-related excitation to the site in the inner ear representing the tone frequency. Thus, with a sufficiently loud noise, the auditory periphery may provide an uninterrupted neuronal excitation at the site representing the tone frequency that could explain perceiving a tone continuing through the noise.

Top-down mechanisms are implicated by the observation that perceptual restoration is affected not only by the stimulus-driven activity in the bottom-up auditory pathway but also by the state of neurons at higher stages in the auditory pathway. In human subjects, for example, the perceptual restoration of words is enhanced if the words are previously known to subjects. Pseudowords that contain the same phonemes as real words but are not found in the native language of the speaker are not as easily restored (Samuel 1996). This indicates that lexical information,

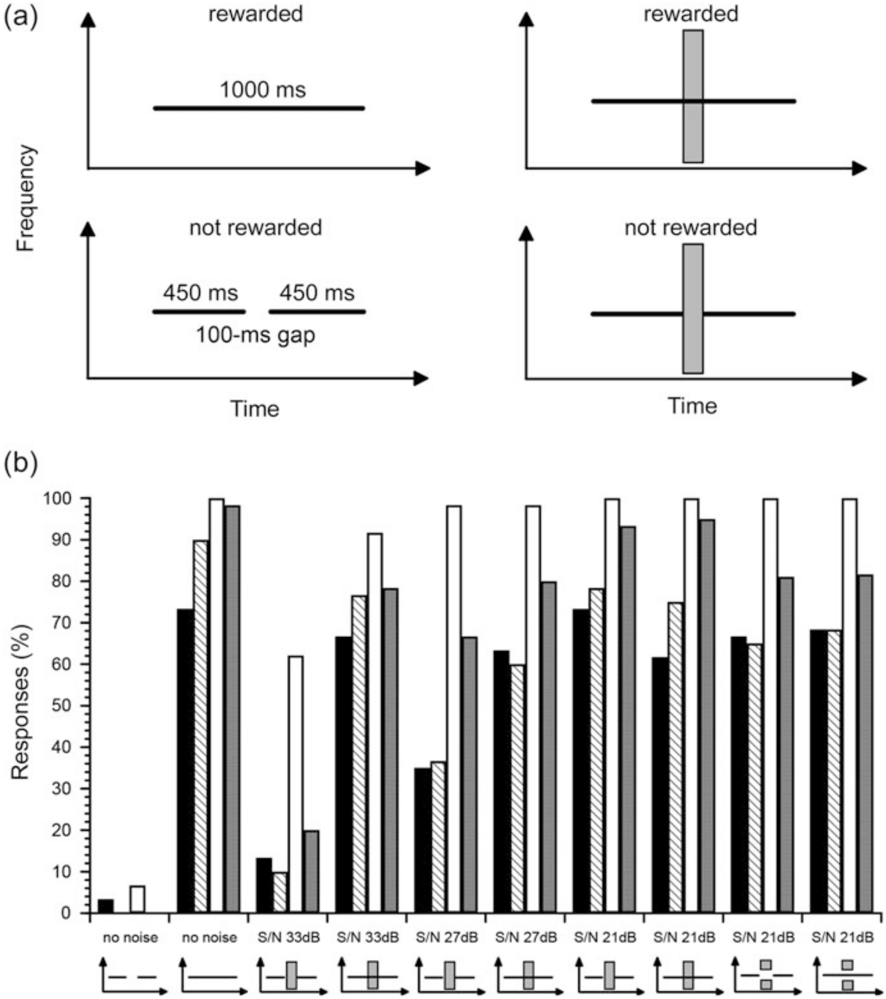


Fig. 3.4 Perceptual restoration in starlings using tonal stimuli. In perceptual restoration, a complete signal (e.g., a tone) can be restored from a signal missing a part, that is, having a silent gap, if the gap is filled with noise exciting the auditory system. Schematic spectrograms of the stimuli used for training starlings for testing perceptual restoration of pure tones are shown in (**a**) together with the reward contingencies. By rewarding only a response to an uninterrupted long tone, the birds were trained to show a go response if presented with the uninterrupted tone and to not respond (no-go response) if the tone had a 100-ms gap in the middle. The probability of a go response of four different individual starlings (differently patterned columns) is shown in (**b**). The small pictograms beneath each set of four columns indicate the stimulus condition together with information on the S/N ratio (i.e., the tone signal level in relation to the spectrum level of the noise). All four birds managed the basic discrimination without added noise (two leftmost sets of columns shown in (**b**)). At a S/N ratio of 33 dB, three out of the four birds only responded weakly to the unrewarded stimulus, while one bird already responded in that condition as if an uninterrupted stimulus had been presented. With decreasing S/N ratio, created by increasing the level of the noise, all birds responded to the interrupted tone presented with the gap filled with noise as if the rewarded uninterrupted tone had been presented. This behavior suggests perceptual restoration. If the S/N ratio was sufficiently small, no spectral energy at the tone frequency was

which is a learned feature of speech, supports the perceptual restoration of phonemes. A corresponding observation has been made in starlings. Braaten and Leary (1999) showed that starlings are better able to perceptually restore a missing starling song segment if it is embedded in a starling song context rather than in a budgerigar (*Melopsittacus undulatus*) song context. This example suggests that the context of species-specific song elements supports perceptual restoration. Using a larger sample of starling song motifs, Seeba and Klump (2009) demonstrated that the perceptual restoration of starling song motifs is better if the song is already familiar to the subjects being tested (Fig. 3.5). Similar to the previous tests on perceptual restoration of tonal signals, silent gaps were introduced into starling song motifs that allowed the birds to distinguish a complete song motif lacking silent gaps (complete motif, Fig. 3.5a) from one interrupted with silent gaps (gap motif, Fig. 3.5b). If the silent gaps were then filled with a noise burst (noise motif, Fig. 3.5c), discrimination from the complete motif was made more difficult. The interpretation of this result is that the birds perceived the song elements removed from the complete motif to be present in the noise motif version of the song but not in the gap motif version. To determine the birds' ability to discriminate the different types of motifs, a repeated background procedure was used. The birds listened to a repeating motif of one type (e.g., a particular complete motif) and were required to report when they heard the repeating sound switch to the same song motif presented as another type of stimulus (e.g., the corresponding noise motif). Their response latency for the discrimination served as an indicator of the saliency of the difference between the two types of the same song motif that had to be discriminated. An analysis of response latencies using multidimensional scaling (MDS) revealed evidence for top-down processing. If individual subjects had prior experience with the song motifs due to the fact that either they or their cage mates sang the motifs (i.e., they had previously experienced the motifs), they showed a better perceptual restoration than if the song motifs were previously unknown (Fig. 3.5b). This pattern of results suggests that in European starlings, both bottom-up and top-down mechanisms contribute to perceptual restoration.

3.6 Auditory Streaming of Sound Sequences

If starlings listen to the songs of conspecifics, which are composed of sequences of motifs that can last up to a minute or more, they need to be able to link together the sequential song motifs belonging to the same song and associate them with one specific source. Moreover, the sequential motifs produced by one starling must be discriminated from the motifs sung by other starlings. Thus, songs must be

---

Fig. 3.4 (continued) necessary to elicit the restoration response (two rightmost sets of columns shown in (**b**); data from a study by Klump et al. 1999)

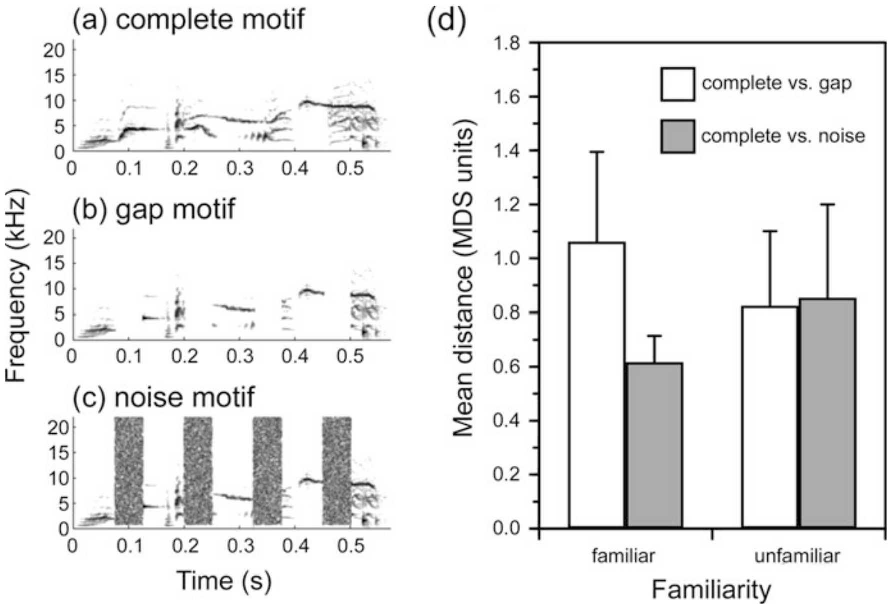


Fig. 3.5 Perceptual restoration in starlings using natural song motifs. From natural songs (**a** complete motif), variants were created that either had silent 50-ms gaps (**b** gap motif) or the same gaps in the signal that were then filled with broadband noise (**c** noise motif). If perceptual restoration is elicited by the noise, then the perceptual distance should be smaller between (**a**) complete motifs and (**c**) noise motifs than between (**a**) complete motifs and (**b**) gap motifs. The data in (**d**) shows that this was observed for song motifs that were familiar to the birds suggesting that perceptual restoration occurred. Mean (+s.d.) perceptual distances were derived from a multidimensional scaling (MDS) measure that was based on response latencies for discriminating the three variants of the song motifs in an operant task. If song motifs were unfamiliar to the birds, the perceptual distance measures did not differ, suggesting that perceptual restoration was less effective in that condition (for details, see Seeba and Klump 2009)

processed in separate auditory streams if a starling aims to evaluate the songs from different, simultaneously singing birds (e.g., a female choosing between males; see Mountjoy and Lemon 1996). This scenario resembles the separate processing by a human listener of sentences produced simultaneously by different talkers.

3.6.1 Subjective and Objective Measures of Auditory Streaming in Humans

In the context of ASA, auditory streaming has been widely studied using the so-called "ABA–" stimulus paradigm first introduced by van Noorden (1975). In this paradigm, A and B denote different tones presented in a long sequence of triplets, with each tone having a specific frequency (or other physical features; see

Sect. 3.6.4). Each tone in a sequence of ABA– triplets can be followed by either another tone or by a silent interval ("–") having a duration that corresponds to the time period from the onset of one tone to that of the next tone in a triplet. Human subjects listening to a sequence of repeating ABA– triplets perceive a galloping rhythm if they process the sequence of tones in one stream (i.e., ABA–ABA–ABA–ABA–. . .). They perceive two parallel isochronous sequences of A and B tones if they process the two sequences in different streams (i.e., A–A–A–A–A. . . and –B– – –B– – – B– – – B . . .). These subjectively different percepts have been used to infer whether listeners are segregating the A and B tones into different streams or whether they group these tones into one stream.

In human psychophysical studies, a "perceptual landscape" has been mapped out by varying not only the frequency difference (Δf) between the A and B tones but also the tone repetition time (TRT). Here, TRT corresponds to the time period from the beginning of the first A tone to that of the B tone in the regular temporal pattern; sometimes TRT is expressed as a percentage relative to the tone duration (TD). Van Noorden (1975) identified two boundaries in the perceptual landscape for ABA– patterns: the fission boundary and the temporal coherence boundary. The fission boundary designates combinations of the Δf between A and B tones and TRT that will result in a nearly obligatory assignment of the two signals to the same stream if smaller values of Δf and larger values of TRT are chosen. In contrast, the temporal coherence boundary designates combinations of Δf and TRT that will result in a nearly obligatory assignment of the A and B tones to separate streams if larger values of Δf and smaller values of TRT are chosen. Between the two boundaries is a region on the map that results either in an ambiguous percept or in a percept in which the A and B tones are initially perceived as one stream, but that become increasingly likely to be perceived as two streams with increasing sequence length. This change in the percept has been termed the "buildup" of stream segregation (Bregman 1978). However, a number of studies have now shown that a two-stream percept can also revert to a one-stream percept, indicating a more or less bistable process in which the interpretation by the brain can assume multiple states (Denham and Winkler 2006; Winkler et al. 2012; Mill et al. 2013).

In contrast to a listener's subjective judgment of whether a particular tone sequence is heard as one or two streams, an objective measure of auditory streaming can be obtained by determining perceptual thresholds that differ depending on whether the sequence is processed as one or two streams (e.g., Micheyl and Oxenham 2010; Dolležal et al. 2014). For example, human listeners have lower thresholds for discriminating differences in temporal patterns when the altered pattern is fully contained within a single stream compared to when the altered pattern spans two segregated streams. If in the ABA– stimulus paradigm, for example, a single B tone is time shifted relative to the positions of its two surrounding A tones, which themselves provide a fixed time pattern for reference, the shift is more easily detected, and shift detection thresholds are lower, if the entire sequence of A and B tones is perceived as one stream with a galloping rhythm.

3.6.2 Subjective and Objective Measures of Auditory Streaming in Starlings

In psychoacoustic studies of auditory streaming using the ABA– paradigm, starlings have been shown to have a similar percept as human listeners. MacDougall-Shackleton et al. (1998) trained starlings to discriminate between constant frequency triplets that were presented either in the sequence AAA–AAA–. . ., which produced a galloping rhythm, or in sequences producing isochronous rhythms (e.g., A–A–A–A–A. . . or –A– – –A– – – A– – – A . . .). The birds were rewarded when they pecked the appropriate "galloping" or "isochronous" key in an operant task. When the birds mastered the discrimination, probe tone sequences of ABA–ABA– . . ., with different A and B tone frequencies, were presented with a low probability, and the choice of response key was observed (no rewards were given). If the A and B tone frequencies were similar and differed by only 0.9 semitones (i.e., less than one half tone in Western tonal music), the birds mostly responded by pecking the "galloping" key, suggesting the A and B tones were integrated into a single stream. However, if the A and B tone frequencies differed by 9 or 21 semitones (where 12 semitones equal one musical octave), the birds mostly responded by pecking the "isochronous" key, suggesting they heard the A and B tones as two different, isochronous rhythms. This pattern of results indicated that the birds have a percept that corresponds well to a human listener's percept of auditory streaming as measured by subjective rhythm perception (Fig. 3.6a, b).

Similar to human listeners, starlings also exhibit higher time-shift detection thresholds if A and B tones have a larger frequency separation (Fig. 3.6c, d). If A and B tones with identical frequency were presented, the starling's time-shift detection threshold was 28 % of the interval (Fig. 3.6d; Itatani and Klump 2014), which resembles the performance of human subjects in a similar task (van Noorden 1975, Micheyl and Oxenham 2010). At larger frequency differences between A and B tones of 6 semitones (1/2 octave) and 12 semitones (1 octave), the starling's time-shift detection threshold was 34 % and 41 %, respectively (Fig. 3.6d). The interpretation of this result is that as the Δf between the A and B tones increased, the likelihood of their segregation into two streams also increased, making the time-shift detection task more difficult. This similarity between starlings and humans in comparing time intervals is not unexpected, since starlings have an ability to discriminate tone duration similar to that of humans (Maier and Klump 1990). In human subjects, objective and subjective measures of auditory streaming are highly correlated with each other (Micheyl and Oxenham 2010; Dolležal et al. 2014). Consequently, objective measures are well suited to evaluate the mechanisms underlying ASA because they are applicable in both psychophysical and neurophysiological studies of stream segregation.

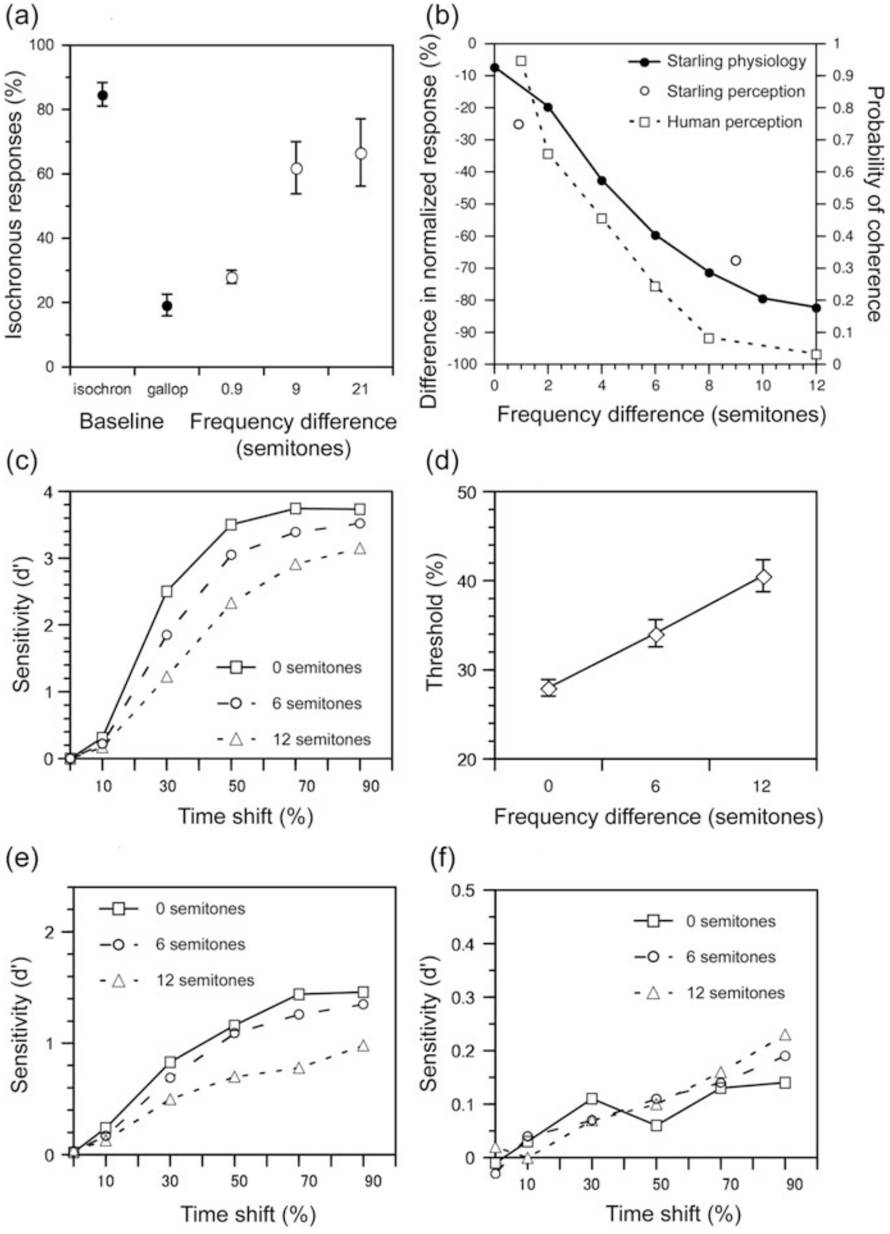


Fig. 3.6 Auditory streaming and its neuronal correlates in starlings. Auditory streaming was studied using a series of repeating ABA– triplet stimuli, where "A" and "B" indicate tones with different frequencies, and "–" indicates an intervening silent interval between triplets in a sequence. If the A and B tones are processed in one stream, a galloping rhythm is perceived; if they are processed in separate streams, two isochronous tone sequences (A and B tones separately) are perceived (van Noorden 1975). (**a**) Starlings could be trained to report in baseline trials their perception of either a galloping rhythm or an isochronous rhythm. On test trials, their assignment of ABA– triplet sequences to one of these two rhythms depended on the frequency difference between A and B tones, with assignments of isochronous being more common at larger frequency

3.6.3 Neural Correlates of Auditory Streaming Based on Frequency Differences

Neurons in the thalamic projection area of the starling auditory forebrain show a response map that in some aspects corresponds well to the perceptual landscape describing auditory streaming in human subjects (Bee and Klump 2004). When starling auditory forebrain neurons were presented with a sequence of repeating ABA– triplets, with the A tone frequency being at the neuron's most sensitive frequency and the B tone frequency differing from that of the A tone, responses to the B tone were reduced compared to A-tone responses (Fig. 3.6b). Reduced responses to the B tone were due, in part, to the frequency tuning of the neurons, but they were also due, in part, to suppression of the B-tone response by the preceding A-tone response. This suppressive effect was especially strong if the A and B tones were temporally abutting (i.e., the TRT was equal to the TD) and if the Δf was small (Bee and Klump 2005). This observation suggests that forward masking supports the segregation of the neuronal responses to the interleaved A and B tones and thus may promote perceptual stream segregation. The difference in the normalized response to the A and B tones in the ABA– triplet pattern relates well to the probability of perceptual stream segregation both in starlings and in human subjects (Fig. 3.6b). Bee and Klump (2004) mapped out a "neural landscape" depicting the difference in normalized responses to A and B tones as functions of Δf and TRT (similar to the perceptual landscape described earlier). In this landscape, iso-response contours paralleled the human fission boundary, further suggesting that the separate neural representation of A and B tones supports perceptual stream segregation.

The neurons in the starling auditory forebrain showed a decreasing response with an ongoing presentation of the repeating ABA– triplets (Bee et al. 2010). This decrease suggests that neural adaptation occurs over a multi-second time period. Such long adaptation times also have been observed in the mammalian auditory

Fig. 3.6 (continued) separations (MacDougall-Shackleton et al. 1998). This observation suggests that starlings have the same subjective streaming percept as humans. (**b**) The difference in the normalized response of starling auditory forebrain neurons to A and B tones (*filled circles*, *solid line*; left *y*-axis) was correlated with starling perception (*open circles*; right *y*-axis) and human perception (*open squares*, *dashed line*; right *y*-axis), both measured as the probability of coherence, which indicates the strength of the one-stream response (see Bee and Klump 2004). (**c**, **d**) The probability for detecting a time shift of one B tone in a sequence of repeating ABA– triplets provides for an objective measure of auditory stream segregation in the starling because it appears easier to detect the time shift within a stream than between tones processed in different streams (Itatani and Klump 2014). The behavioral sensitivity for detecting the time shift and the detection threshold (criterion $d' = 1$) are shown in (**c**) and (**d**), respectively, in relation to the frequency difference between A and B tones. (**e**) The sensitivity derived from the temporal pattern in the response of starling forebrain neurons parallels the behaviorally determined sensitivity depicted in panel (**c**), whereas no correspondence between a response rate-based sensitivity measure (**f**) and the behavioral response was found (Itatani and Klump 2014) cortex (e.g., Ulanovsky et al. 2004; Asari and Zador 2009; von der Behrens et al. 2009). The magnitude of the decrease was generally larger at short TRTs, for which the rate of adaptation was also faster. Furthermore, at a smaller Δf the B-tone response adapted more than at larger values of Δf. Taken together, these effects could be used to model neurometric functions that predict how the probability of a two-stream percept develops over time (Bee et al. 2010). Neurometric functions computed on the basis of the starling's neural responses to sequences of ABA– triplets suggest that this bird species should experience a buildup of stream segregation, like human subjects do. Unfortunately, behavioral data are still lacking from starlings or other nonhuman species that could be used to investigate the correlation between adaptation of neuronal responses over time and the perceptual segregation of auditory streams.

Until recently, all neuronal measures of auditory streaming recorded in animal models at the cellular level were obtained from either anesthetized (Pressnitzer et al. 2008) or awake (Bee and Klump 2004, 2005; Fishman et al. 2001, 2004; Micheyl et al. 2005) animals that listened passively to stimuli. By applying an objective behavioral measure of auditory stream segregation, Itatani and Klump (2014) were able, for the first time, to investigate neuronal responses in the auditory forebrain of an actively listening animal, while it was indicating its current perceptual state regarding stream segregation. Itatani and Klump (2014) presented sequences of ABA– triplets to starlings while recording from an area in the starling brain that is the homologue of the thalamic projection layer in the mammalian primary auditory cortex, as in previous studies (e.g., Bee and Klump 2004, 2005; Bee et al. 2010). While recording the neurons' responses, however, they engaged the birds in the tone-shift detection task described above, which both humans and starlings find more difficult to solve when A and B tones are processed as separate streams. Specifically, the birds had to detect a single time-shifted B tone in an ongoing series of ABA– triplets in which all other B tones were perfectly centered in time with respect to the leading and trailing A tones. As described above, the starling's sensitivity d' for detecting the time shift was related to the Δf between the A and B tones (Fig. 3.6c). In general, it appears to be more difficult to assess temporal relationships between signal elements if they are represented by separate analysis channels of the auditory system, which promotes the processing of segregated auditory streams [e.g., as proposed by the channeling hypothesis of streaming proposed by Hartmann and Johnson (1991)].

Applying signal detection theory, Itatani and Klump (2014) determined two neuronal response measures – one based on spike timing and the second based on spike rate – and compared them using a ROC analysis that determined the sensitivity d_a (i.e., a measure corresponding to d' that is derived from analyzing the area under the ROC curve) for discriminating the time-shifted B tone relative to a regularly timed tone. The first response measure (Fig. 3.6e) used a van Rossum (2001) analysis to compare the temporal patterning of the spike train during the ABA– triplet with the time-shifted B tone to the corresponding response elicited by the B tone in the regular and immediately preceding ABA– triplet. In this analysis, the time windows for determining B-tone responses extended for a time period of

80 ms and started 80 ms after the onset of the relevant ABA– triplet; this ensured the window always included the B-tone responses. The response measure that was subjected to an ROC analysis was the van Rossum distance obtained with a time constant of 3.16 or 31.6 ms, which represent coincident spikes and the ongoing spiking rate, respectively. The second response measure (Fig. 3.6f) was simply based on the neurons' rate response in a temporal window that was adjusted to the time of presentation of the B tone (i.e., having a start time that accounted for the latency of the neural response and the time shift of the B tone). Each temporal window had a duration of 15 or 40 ms that captured the onset and ongoing portions of the neural response, respectively. For the ROC analysis, this rate response was compared to the response to the regularly timed B tone in the ABA– triplet immediately before the triplet with the time-shifted B tone. In comparing results from these two different measures of neuronal responsiveness, the analysis based on spike timing (Fig. 3.6e) revealed a much higher neuronal sensitivity for detecting the time shift than the corresponding analysis based on spike rates (Fig. 3.6f). The temporal analysis was also much better at recovering the effects of Δf on tone detection. The highest measures of sensitivity based on these analyses of the responses of small clusters of neurons reached only half the maximum value of the corresponding behavioral sensitivity measure (cf. Fig. 3.6c, e). Nevertheless, it seems likely that the behavioral response thresholds can be explained by the neuronal temporal response patterns given that the neural recordings reflect the operation of only a few cells near the electrode tip, whereas the behavioral detection task likely engages a much larger population of neurons.

Based on their objective measures of auditory streaming, Itatani and Klump (2014) also investigated whether the starling primary auditory cortex, in addition to representing features of the stimulus, also represents the decision taken by the bird when discriminating the sounds. To this end, they compared the responses in trials in which the birds correctly detected the time shift (hits) to those trials in which the birds did not respond when there was a time shift (misses). Interestingly, there was no difference between the responses in these two conditions. This result, which indicates that this cortical area represents stimulus features used for auditory streaming but not the decision of the animal, corresponds well to observations made in human subjects (Gutschalk et al. 2008; Dykstra et al. 2011).

3.6.4 Auditory Streaming Based on Temporal Differences

Studies of streaming in human listeners have shown that a large number of features, in addition to frequency differences, promote stream segregation, provided that differences in the features between the signals to be segregated are perceptually quite salient (Moore and Gockel 2002, 2012). For example, a number of temporal features of sound can promote auditory streaming in humans. Grimault et al. (2002) demonstrated that sinusoidally amplitude-modulated noise bursts presented as a sequence of ABA– triplets (with A and B denoting signals differing in modulation frequency but having the same carrier signal) will be perceived as segregated into different auditory streams if the A and B signals differ in the rate of amplitude modulation. Dolležal et al. (2012) found similar results presenting sinusoidally amplitude-modulated tones in ABA– triplet sequences to human subjects. For conditions in which all three frequency components of the amplitude-modulated tones were confined to an auditory filter, stream segregation was observed if the modulation frequencies differed by about 0.6–0.8 octaves. This observation is similar to the observation by Grimault et al. (2002), who reported clear stream segregation for a difference in modulation frequency of at least 0.9 octaves. Both studies also observed that stream segregation is reduced by a decrease in modulation depth. Dolležal et al. (2014) compared subjective and objective measures of stream segregation using the sinusoidally amplitude-modulated tones and correlated these with measures of the blood oxygenation level-dependent (BOLD) response in human subjects indicating the activation of specific brain areas in fMRI. As was observed for the streaming of pure tone series, objective (i.e., time-shift detection thresholds) and subjective measures of stream segregation were correlated. However, there was only a weak correlation between the BOLD response and the degree of stream segregation across different conditions representing different carrier frequencies and modulation frequencies. An alternate manipulation of stimuli that only modifies the temporal pattern but not the spectral composition builds on changing the phase relation between frequency components in a harmonic complex. Using the ABA– paradigm, Roberts et al. (2002) and Dolležal et al. (2012) demonstrated harmonic tone complexes can be processed in different streams if one of the complexes has all components in cosine phase (e.g., the A signal) and the other complex has the components in random phase (e.g., the B signal). Stream segregation was also observed using harmonic complexes having components in cosine phase and complexes having components in alternating phase (Roberts et al. 2002).

Birdsong also contains rapid amplitude modulations, and song elements are sometimes composed of harmonic tone complexes (Chap. 2). Therefore, we should expect that a bird like the European starling possesses mechanisms for processing amplitude modulations and for evaluating the temporal fine structure of harmonic complex sounds in auditory streaming. Similar to the approach taken with pure tones (Bee and Klump 2004), Itatani and Klump (2009, 2011) evaluated whether temporal cues were represented by separate populations of neurons based on different modulation rates of tones having the same carrier frequency (Itatani and Klump 2009) or by different temporal fine structures, which were created by manipulating the phase relationships of components in harmonic complexes (Itatani and Klump 2011). Furthermore, Itatani and Klump (2009, 2011) investigated whether suppression was involved in supporting the neural representation of the signals in different streams by different populations of neurons, similar to the neuronal mechanism underlying the streaming of pure tones. As was found for the streaming of pure tone sequences, starling auditory forebrain neurons showed a lower response rate for temporally abutting A and B signals differing in temporal cues than for A and B signals that were well separated in time (offset-to-onset interval between A and B signals of at least 300 % of the signal duration). The effect of introducing silent intervals between the signals, however, was independent of the effects that the temporal structure of the stimuli had on the neuronal response rate. This was the case for manipulations of both amplitude modulation and temporal fine structure. These results, together with those of Bee and Klump (2004, 2005), suggest that the forward masking effects provided by the peripheral pattern of excitation related to the spectral tuning of the signals were independent from suppressive effects related to the temporal structure of the signals, for which both effects were additive. The differences in the rate responses of auditory forebrain neurons that were elicited by manipulations of temporal fine structure were as expected based on differences of excitation in the auditory periphery (random phase complex tone stimuli possibly being subject to less compression than, e.g., cosine phase complex tone stimuli; see Stainsby et al. 2004). Since the different recording sites in the starling forebrain responded preferentially to specific temporal features of the sound (either modulation frequency or type of phase relation between components of harmonic tone complexes), it can be inferred that separate populations of neurons may represent the sounds associated with the separate streams, as was the case for pure tone stimuli differing in frequency.

So far, we have focused on the rate response of forebrain neurons in representing amplitude-modulated tones and harmonic complex tones with different phase relations. However, temporal features of a sound can also be represented by temporal patterns in the response (Chap. 2). Itatani and Klump (2009) estimated the vector strength of the neuronal response associated with the period histogram in relation to the period of the envelope modulation imposed on modulated tones. They found significant effects of the modulation frequency difference on the response measure for both the rate and the temporal response measure (i.e., the vector strength of the response related to the period of the sinusoidal envelope modulation). Both the rate and the temporal response measures would have been suitable to produce separate representations in the brain activity pattern allowing for auditory stream segregation. In their study on the possibility of stream segregation by differences in the phase relation between components of a harmonic tone complex, Itatani and Klump (2011) quantified the magnitude of phase locking to the fundamental of incoming stimuli by evaluating the peak height at the interval corresponding to the *f*0 (fundamental frequency) of the stimuli in the all order inter-spike-interval histogram. The type of phase relation between the components showed a somewhat smaller effect on the peak-to-background ratio (i.e., the relative magnitude of the *f*0 peak normalized by the magnitude of the background in the inter-spike-interval histogram) than on the neurons' rate of spiking. Thus, for this type of stimulus, the temporal response was less suited than the rate response for differentiating between the signals represented in separate streams. Furthermore, the neurons' temporal response measure in the streaming of harmonic complex tones with different phase relations is much more affected by the signal duration than is the rate response measure (Dolležal et al. 2012). Reducing the signal duration from 125 to 40 ms deteriorates the temporal response measure, whereas the rate response is less affected. In addition, the temporal response measure cannot represent well the effect of the phase relation for an *f*0 of 400 Hz. For such a high *f*0, the neurons in the starling forebrain do not fire time-locked action potentials to the period of the signal envelope (2.5 ms) and, thus, cannot represent the temporal structure by a temporal response. Both observations of the neuronal response suggest a better match between the rate response and the streaming percept than between the temporal response measure and the percept.

3.7 Summary and Future Directions

The research on European starlings reviewed in this chapter highlights several key aspects of the perceptual organization of acoustic scenes in the context of animal communication. First, animal vocalizations, including human speech and birdsong, are often comprised of complex sequences of elements with spectro-temporally complex acoustic features. In starlings, songs are learned and play key roles in various behavioral and social contexts, including mate choice and individual recognition. Second, communication using complex vocal sequences often takes place in large, noisy social aggregations, such as a human cocktail party or an animal breeding colony. Important aspects of social communication by starlings take place in the cacophonous environments of a dawn chorus or a large night roost. Third, the demands of vocal communication in a crowd select for the evolution of mechanisms responsible for segregating signals of interest from other concurrent signals and from the din of background noise created in the aggregation, as well as for perceptually binding together signal elements produced by a common source. Current evidence from studies of CMR, CDD, perceptual restoration, and auditory streaming indicates that starlings likely exploit many of the same underlying cues that humans use in perceptually organizing complex acoustic scenes (e.g., differences in frequency, differences in patterns of amplitude modulation, differences in the temporal fine structure of harmonic sounds). Parallel work on ASA in other animals, such as frogs (Bee 2012, 2015; Chap. 4) and insects (Römer 2013), suggests that a common set of physical features of sounds are likely processed using bottom-up mechanisms across a wide range of animals. The extent to which these mechanisms are evolutionarily homologous across taxa or represent instances of convergent evolution remains an important and open question in comparative neurobiology. Fourth, the integration of psychophysical and neurophysiological tools, and in particular the use of objective measures of perceptual performance, provides a basis for linking complex perception to its underlying neural mechanisms at a cellular level. In this regard, starlings represent one of the best-studied animal models of ASA. Results from this work have found several neural correlates of perception and identified underlying neural codes for ASA, for example, in terms of describing the relative importance of rate codes versus temporal codes in various listening tasks.

Although we find many neuronal correlates of ASA in starlings, the causality of auditory object formation is still unclear. Recent studies, for example, have provided evidence for the integration of multiple cues in auditory stream formation. Multiple cues have been demonstrated to both support the grouping of sounds forming auditory streams or objects (e.g., harmonic relationships between components, common onset, common envelope variation) and to promote the segregation of sounds and their assignment to different streams or objects (e.g., spectral frequency differences, Bregman 1990). The weight given to the different cues supporting integration versus segregation will determine which percepts prevail, that is, whether the auditory system engages in grouping or segregation. Elhilali et al. (2009) have demonstrated that common sound onset (or more generally, temporal coherence of sounds; see review by Shamma and Micheyl 2010) can be such a strong grouping cue that it will override other cues that otherwise would have been sufficient for segregating streams if presented alone (such as frequency differences large enough to perceive overlapping sounds as separate streams). It is in resolving these various outstanding questions where future studies of ASA in the context of animal communication could have their highest payoff. If we could discover how the mechanisms for grouping and segregation resulting from the rich set of cues provided by natural stimuli, such as communication signals, are integrated, we will get closer to an understanding of how perception of auditory objects or streams in real-life situations is achieved. Given that starlings show many parallels to humans in perceptual tasks related to ASA, they may also provide the key to understanding how neuronal mechanisms function in enabling them to learn new songs in the cacophony of a roost or to find the best mate in the singer's contest in the breeding colony.

Acknowledgments The starling studies were funded by grants from the Deutsche Forschungsgemeinschaft (SFB/TRR31, FOR 306, GRK 591) and by a grant from the National Science Foundation (INT-0107304, postdoctoral fellowships to M. A. Bee). I thank the editors for their feedback on a previous version of the chapter.

References

Adret-Hausberger M, Jenkins PF (1988) Complex organization of the warbling song in the European starling *Sturnus vulgaris*. Behavior 107:138–156. doi:10.1163/156853988X00322

Asari H, Zador AM (2009) Long-lasting context dependence constrains neural encoding models in rodent auditory cortex. J Neurophysiol 102:2638–2656. doi:10.1152/jn.00577.2009

Baugh AT, Ryan MJ, Bernal XE, Rand AS, Bee MA (2016) Female túngara frogs do not experience the continuity illusion. Behav Neurosci 130:62–74. doi:10.1037/bne0000115

Bee MA (2012) Sound source perception in anuran amphibians. Curr Op Neurobiol 22:301–310. doi:10.1016/j.conb.2011.12.014

Bee MA (2015) Treefrogs as animal models for research on auditory scene analysis and the cocktail party problem. Int J Psychophysiol 95:216–237. doi:10.1016/j.ijpsycho.2014.01.004

Bee MA, Klump GM (2004) Primitive auditory stream segregation: a neurophysiological study in the songbird forebrain. J Neurophysiol 92:1088–1104. doi:10.1152/jn.00884.200

Bee MA, Klump GM (2005) Auditory stream segregation in the songbird forebrain: effects of time intervals on responses to interleaved tone sequences. Brain Behav Evol 66:197–214. doi:10.1159/000087854

Bee MA, Micheyl C (2008) The “cocktail party problem”: what is it? How can it be solved? And why should animal behaviorists study it? J Comp Psychol 122:235–251. doi:10.1037/0735-7036.122.3.235

Bee MA, Buschermöhle M, Klump GM (2007) Detecting modulated signals in modulated noise: (II) neural thresholds in the songbird forebrain. Eur J Neurosci 26:1979–1994. doi:10.1111/j.1460-9568.2007.05805.x

Bee MA, Micheyl C, Oxenham AJ, Klump GM (2010) Neural adaptation to tone sequences in the songbird forebrain: patterns, determinants, and relation to the build-up of auditory streaming. J Comp Physiol A 196:543–557. doi:10.1007/s00359-010-0542-4

Bigalke-Kunz B, Rübsamen R, Dörrscheidt GJ (1987) Tonotopic organization and functional characterization of the auditory thalamus in a songbird, the European starling. J Comp Physiol A 161:255–265. doi:10.1007/BF00615245

Borrill SJ, Moore BC (2002) Evidence that comodulation detection differences depend on within-channel mechanisms. J Acoust Soc Am 111:309–319. doi:10.1121/1.1426373

Braaten RF, Leary JC (1999) Temporal induction of missing birdsong segments in European Starling. Psychol Sci 10:162–166. doi:10.1111/1467-9280.00125

Bregman AS (1978) Auditory streaming is cumulative. J Exp Psychol Hum Percept Perform 4:380–387. doi:10.1037/0096-1523.4.3.380

Bregman AS (1990) Auditory scene analysis. MIT Press, Cambridge

Bregman MR, Patel AD, Gentner TQ (2016) Songbirds use spectral shape, not pitch, for sound pattern recognition. Proc Natl Acad Sci USA 113:1666–1671. doi:10.1073/pnas.1515380113

Buschermöhle M, Verhey JL, Feudel U, Freund JA (2007) The role of the auditory periphery in comodulation detection difference and comodulation masking release. Biol Cybern 97:397–411. doi:10.1007/s00422-007-0179-8

Carr CE, Code RA (2000) The central auditory system of reptiles and birds. In: Dooling RJ, Fay RR, Popper AN (eds) Comparative hearing: birds and reptiles. Springer, New York, pp. 197–248

Chaiken M, Böhner J, Marler P (1993) Song acquisition in European starlings, *Sturnus vulgaris*: a comparison of the songs of live-tutored, tape-tutored, untutored, and wild-caught males. Anim Behav 46:1079–1090. doi:10.1006/anbe.1993.1298

Chaiken M, Böhner J, Marler P (1994) Repertoire turnover and the timing of song acquisition in European starlings. Behaviour 128:25–39. doi:10.1163/156853994X00037

Comins JA, Gentner TQ (2014) Temporal pattern processing in songbirds. Curr Opin Neurobiol 28:179–187. doi:10.1016/j.conb.2014.08.003

Denham SL, Winkler I (2006) The role of predictive models in the formation of auditory streams. J Physiol Paris 100:154–170. doi:10.1016/j.jphysparis.2006.09.012

Dolležal L-V, Itatani N, Günther S, Klump GM (2012) Auditory streaming by phase relations between components of harmonic complexes: a comparative study of human subjects and bird forebrain neurons. Behav Neurosci 126:797–808. doi:10.1037/a0030249

Dolležal L-V, Brechmann A, Klump GM, Deike S (2014) Evaluating auditory stream segregation of SAM tone sequences by subjective and objective psychoacoustical tasks, and brain activity. Front Neurosci 8:119. doi:10.3389/fnins.2014.00119

Dykstra AR, Halgren E, Thesen T, Carlson CE, Doyle W, Madsen JR, Eskandar EN, Cash SS (2011) Widespread brain areas engaged during a classical auditory streaming task revealed by intracranial EEG. Front Hum Neurosci 5:74. doi:10.3389/fnhum.2011.00074

Eens M, Pinxten R, Verheyen RF (1989) Temporal and sequential organization of song bouts in the starling. Ardea 77:75–86

Eens M, Pinxten R, Verheyen RF (1991) Male song as a cue for mate choice in the European starling. Behaviour 116:210–238. doi:10.1163/156853991X00049

Eens M, Pinxten R, Verheyen RF (1992a) Song learning in captive European starlings, *Sturnus vulgaris*. Anim Behav 44:1131–1143. doi:10.1016/S0003-3472(05)80325-2

Eens M, Pinxten R, Verheyen RF (1992b) No overlap in song repertoire size between yearling and older starlings *Sturnus vulgaris*. Ibis 134:72–76. doi:10.1111/j.1474-919X.1992.tb07233.x

Elhilali M, Ma L, Micheyl C, Oxenham AJ, Shamma SA (2009) Temporal coherence in the perceptual organization and cortical representation of auditory scenes. Neuron 61:317–329. doi:10.1016/j.neuron.2008.12.005

Ficken RS, Ficken MS, Hailman JP (1974) Temporal pattern shifts to avoid acoustic interference in singing birds. Science 183:762–763. doi:10.1126/science.183.4126.762

Fishman YI, Reser DH, Arezzo JC, Steinschneider M (2001) Neural correlates of auditory stream segregation in primary auditory cortex of the awake monkey. Hear Res 151:167–187. doi:10.1016/S0378-5955(00)00224-0

Fishman YI, Arezzo JC, Steinschneider M (2004) Auditory stream segregation in monkey auditory cortex: effects of frequency separation, presentation rate, and tone duration. J Acoust Soc Am 116:1656–1670. doi:10.1121/1.1778903

Gentner TQ, Hulse SH (2000) Female European starling preference and choice for variation in conspecific male song. Anim Behav 59:443–458. doi:10.1006/anbe.1999.1313

Gentner TQ, Margoliash D (2003) Neuronal populations and single cells representing learned auditory objects. Nature 424:669–674. doi:10.1038/nature01731

George I, Cousillas H (2013) How social experience shapes song representation in the brain of starlings. J Physiol Paris 107:170–177. doi:10.1016/j.jphysparis.2012.12.002

Gleich O (1989) Auditory primary afferents in the starling – correlation of function and morphology. Hear Res 37:255–267. doi:10.1016/0378-5955(89)90026-9

Gleich O, Klump GM (1995) Temporal modulation transfer functions in the European Starling (*Sturnus vulgaris*): II. Responses of auditory-nerve fibres. Hear Res 82:81–92. doi:10.1016/0378-5955(94)00168-P

Gleich O, Manley GA (1988) Quantitative morphological analysis of the sensory epithelium of the starling and pigeon basilar papilla. Hear Res 34:69–85. doi:10.1016/0378-5955(88)90052-4

Gleich O, Narins PM (1988) The phase response of primary auditory afferents in a songbird (*Sturnus vulgaris* L.). Hear Res 32:81–91. doi:10.1016/0378-5955(88)90148-7

Grimault N, Bacon SP, Micheyl C (2002) Auditory stream segregation onthe basis of amplitude-modulation rate. J Acoust Soc Am 111:1340–1348

Gutschalk A, Micheyl C, Oxenham AJ, Wilson EC, Melcher JR (2008) Neural correlates of auditory perceptual awareness under informational masking. PLoS Biol 6(6):e138. doi:10.1371/journal.pbio.0060138

Hall JW, Haggard MP, Fernandes MA (1984) Detection in noise by spectro-temporal pattern analysis. J Acoust Soc Am 76:50–56. doi:10.1121/1.391005

Hartmann WM, Johnson D (1991) Stream segregation and peripheral channeling. Music Percept 9:155–184. doi:10.2307/40285527

Hausberger M, Bigot E, Clergeau P (2008) Dialect use in large assemblies: a study in European starling *Sturnus vulgaris* roosts. J Avian Biol 39:672–682. doi:10.1111/j.1600-048X.2008.04307.x

Hofer SB, Klump GM (2003) Within- and across-channel processing in auditory masking: a physiological study in the songbird forebrain. J Neurosci 23:5732–5739. doi:10.1037/a0030185

Hose B, Langner G, Scheich H (1987) Topographic representation of periodicities in the forebrain of the mynah bird: one map for pitch and rhythm? Brain Res 422:367–373. doi:10.1016/0006-8993(87)90946-2

Hulse SH (2002) Auditory scene analysis in animal communication. In: Slater PJB, Rosenblatt JS, Snowdon CT et al (eds) Advances in the study of behavior, vol 31. Academic, San Diego, pp 163–200. doi:10.1016/S0065-3454(02)80008-0

Itatani N, Klump GM (2009) Auditory streaming of amplitude-modulated sounds in the songbird forebrain. J Neurophysiol 101:3212–3225. doi:10.1152/jn.91333.2008

Itatani N, Klump GM (2011) Neural correlates of auditory streaming of harmonic complex sounds with different phase relations in the songbird forebrain. J Neurophysiol 105:188–199. doi:10.1152/jn.00496.2010

Itatani N, Klump GM (2014) Neural correlates of auditory streaming in an objective behavioral task. Proc Natl Acad Sci USA 111:10738–10743. doi:10.1073/pnas.1321487111

Jeanne JM, Thompson JV, Sharpee TO, Gentner TQ (2011) Emergence of learned categorical representations within an auditory forebrain circuit. J Neurosci 31:2595–2606. doi:10.1523/jneurosci.3930-10.2011

Khouri L, Nelken I (2015) Detecting the unexpected. Curr Opin Neurobiol 35:142–147. doi:10.1016/j.conb.2015.08.003

Klump GM, Langemann U (1995) Comodulation masking release in a songbird. Hear Res 87:157–164. doi:10.1016/0378-5955(95)00087-K

Klump GM, Larsen ON (1992) Azimuth sound localization in the European starling (*Sturnus vulgaris*): I. Physical binaural cues. J Comp Physiol A 170:243–251

Klump GM, Nieder A (2001) Release from masking in fluctuating background noise is represented in a songbird's auditory forebrain. Neuroreport 12:1825–1829. doi:10.1046/j.0953-816x.2001.01465.x

Klump GM, Fichtel C, Hamann I, Langemann U (1999) Filling in the gap: evidence for apparent continuity in the songbird auditory system. ARO Midwinter Research Meeting, Abstract 108

Klump GM, Langemann U, Friebe A, Hamann I (2001) An animal model for studying across-channel processes: CMR and MDI in the European starling. In: Breebart DJ, Houtsma AJM, Kohlrausch A, Prijs VF, Schoonhoven R (eds) Physiological and psychophysical bases of auditory function. Shaker Publishing BV, Maastricht, pp. 266–272

Knipschild M, Dörrscheidt GJ, Rübsamen R (1992) Setting complex tasks to single units in the avian auditory forebrain. I: Processing of complex artificial stimuli. Hear Res 57:216–230. doi:10.1016/0378-5955(92)90153-E

Knudsen DP, Gentner TQ (2013) Active recognition enhances the representation of behaviorally relevant information in single auditory forebrain neurons. J Neurophysiol 109:1690–1703. doi:10.1152/jn.00461.2012

Kobayasi KI, Usami A, Riquimaroux H (2012) Behavioral evidence for auditory induction in a species of rodent: Mongolian gerbil (*Meriones unguiculatus*). J Acoust Soc Amer 132:4063–4068. doi:10.1121/1.4763546

Langemann U, Klump GM (2001) Signal detection in amplitude-modulated maskers: I. Behavioral auditory threshold in a songbird. Eur J Neurosci 13:1025–1032. doi:10.1046/j.0953-816x.2001.01464.x

Langemann U, Klump GM (2007) Detecting modulated signals in modulated noise: (I) behavioural auditory thresholds in a songbird. Eur J Neurosci 26:1969–1978. doi:10.1111/j.1460-9568.2007.05804.x

MacDougall-Shackleton SA, Hulse SH, Gentner TQ, White W (1998) Auditory scene analysis by European starlings (*Sturnus vulgaris*): perceptual segregation of tone sequences. J Acoust Soc Am 103:3581–3587. doi:10.1121/1.423063

Maier EH, Klump GM (1990) Auditory duration discrimination in the European starling (*Sturnus vulgaris*). J Acoust Soc Am 88:616–621. doi:10.1121/1.399765

McFadden D (1987) Comodulation detection differences using noise-band signals. J Acoust Soc Am 81:1519–1527. doi:10.1121/1.394504

Micheyl C, Oxenham AJ (2010) Objective and subjective psychophysical measures of auditory stream integration and segregation. J Assoc Res Otolaryngol 11:709–724. doi:10.1007/s10162-010-0227-2

Micheyl C, Tian B, Carlyon RP, Rauschecker JP (2005) Perceptual organization of tone sequences in the auditory cortex of awake Macaques. Neuron 48:139–148. doi:10.1016/j.neuron.2005.08.039

Mill RW, Bőhm TM, Bendixen A, Winkler I, Denham SL (2013) Modelling the emergence and dynamics of perceptual organisation in auditory streaming. PLoS Comput Biol 9:e1002925. doi:10.1371/journal.pcbi.1002925

Miller CT, Dibble E, Hauser MD (2001) Amodal completion of acoustic signals by a nonhuman primate. Nat Neurosci 4:783–784. doi:10.1038/90481

Moore BCJ, Gockel H (2002) Factors influencing sequential stream segregation. Acta Acust/Acust 88:320–333

Moore BCJ, Gockel H (2012) Properties of auditory stream formation. Phil Trans R Soc B 367:919–931. doi:10.1098/rstb.2011.0355

Mountjoy DJ, Lemon RE (1996) Female choice for complex song in the European starling: a field experiment. Behav Ecol Sociobiol 38:65–71. doi:10.1007/s002650050218

Näätänen R, Jacobsen T, Winkler I (2005) Memory-based or afferent processes in mismatch negativity (MMN): a review of the evidence. Psychophysiology 42:25–32. doi:10.1111/j.1469-8986.2005.00256.x

Nelken I, Rotman Y, Bar Yosef O (1999) Responses of auditory-cortex neurons to structural features of natural sounds. Nature 397:154–157. doi:10.1038/16456

Nieder A, Klump GM (2001) Signal detection in amplitude-modulated maskers: II. Processing in the songbird's auditory forebrain. Eur J Neurosci 13:1033–1044

Petkov C, Sutter M (2011) Evolutionary conservation and neuronal mechanisms of auditory perceptual restoration. Hear Res 271:54–65. doi:10.1016/j.heares.2010.05.011

Petkov CI, O'Connor KN, Sutter ML (2003) Illusory sound perception in macaque monkeys. J Neurosci 23:9155–9161

Petkov CI, O'Connor KN, Sutter ML (2007) Encoding of illusory continuity in primary auditory cortex. Neuron 54:153–165. doi:10.1016/j.neuron.2007.02.031

Pressnitzer D, Sayles M, Micheyl C, Winter IM (2008) Perceptual organization of sound begins in the auditory periphery. Curr Biol 18:1124–1128. doi:10.1016/j.cub.2008.06.053

Roberts B, Glasberg BR, Moore BCJ (2002) Primitive stream segregation of tone sequences without differences in fundamental frequency or passband. J Acoust Soc Am 112:2074–2085. doi:10.1121/1.1508784

Römer H (2013) Masking by noise in acoustic insects: problems and solutions. In: Brumm H (ed) Animal communication and noise. Animal signals and communication. Springer, New York, pp 33–63. doi:10.1007/978-3-642-41494-7_3

Rübsamen R, Dörrscheidt GJ (1986) Tonotopic organization of the auditory forebrain in a songbird, the European starling. J Comp Physiol A 158:639–664. doi:10.1007/BF00603820

Samuel AG (1996) Does lexical information influence the perceptual restoration of phonemes? J Exp Psychol Gen 125:28–51. doi:10.1037/0096-3445.125.1.28

Schooneveldt GP, Moore BC (1987) Comodulation masking release (CMR): effects of signal frequency, flanking-band frequency, masker bandwidth, flanking-band level, and monotic versus dichotic presentation of the flanking band. J Acoust Soc Am 82:1944–1956. doi:10.1121/1.395639

Schooneveldt GP, Moore BC (1989) Comodulation masking release (CMR) as a function of masker bandwidth, modulator bandwidth, and signal duration. J Acoust Soc Am 85:273–281. doi:10.1121/1.397734

Schwartz JJ, Huth K, Jones SH, Brown R, Marks J, Yang XL (2010) Tests for call restoration in the gray treefrog Hyla versicolor. Bioacoustics 20:59–86. doi:10.1080/09524622.2011.9753633

Seeba F, Klump GM (2009) Stimulus familiarity affects perceptual restoration in the European starling (*Sturnus vulgaris*). PLoS One 4:e5974. doi:10.1371/journal.pone.0005974

Seeba F, Schwartz JJ, Bee MA (2010) Testing an auditory illusion in frogs: perceptual restoration or sensory bias? Anim Behav 79:1317–1328. doi:10.1016/j.anbehav.2010.03.004

Shamma SA, Micheyl C (2010) Behind the scenes of auditory perception. Curr Opin Neurobiol 20:361–366. doi:10.1016/j.conb.2010.03.009

Singh NC, Theunissen FE (2003) Modulation spectra of natural sounds and ethological theories of auditory processing. J Acoust Soc Am 114:3394–3411. doi:10.1121/1.1624067

Soha JA, Peters S (2015) Vocal learning in songbirds and humans: a retrospective in honor of Peter Marler. Ethology 121:933–945. doi:10.1111/eth.12415

Stainsby TH, Moore BC, Medland PJ, Glasberg BR (2004) Sequential streaming and effective level differences due to phase-spectrum manipulations. J Acoust Soc Am 115:1665–1673. doi:10.1121/1.1650288

Sugita Y (1997) Neuronal correlates of auditory induction in the cat cortex. Neuroreport 8:1155–1159

Theunissen FE, Elie JE (2014) Neural processing of natural sounds. Nat Rev Neurosci 15:355–366. doi:10.1038/nrn3731

Thompson JV, Gentner TQ (2010) Song recognition learning and stimulus-specific weakening of neural responses in the avian auditory forebrain. J Neurophysiol 103:1785–1797. doi:10.1152/jn.00885.2009

Todt D, Naguib M (2000) Vocal interactions in birds: the use of song as a model in communication. In: Slater PJB, Rosenblatt JS, Snowdon CT; et al (eds) Advances in the study of behavior, vol 29. Academic, San Diego, pp 247–296. doi:10.1016/S0065-3454(08)60107-2

Ulanovsky N, Las L, Farkas D, Nelken I (2004) Multiple time scales of adaptation in auditory cortex neurons. J Neurosci 24:10440–10453. doi:10.1523/jneurosci.1905-04.2004

van Noorden LPAS (1975) Temporal coherence in the perception of tone sequences. PhD Dissertation, Technical University Eindhoven

van Rossum MW (2001) A novel spike distance. Neur Compu 13:751–763. doi:10.1162/089976601300014321

Vélez A, Bee MA (2011) Dip listening and the cocktail party problem in grey treefrogs: signal recognition in temporally fluctuating noise. Anim Behav 82:1319–1327. doi:10.1016/j.anbehav.2011.09.015

Verhey JL, Pressnitzer D, Winter IM (2003) The psychophysics and physiology of comodulation masking release. Exp Brain Res 153:405–417. doi:10.1007/s00221-003-1607-1

von der Behrens W, Bauerle P, Kossl M, Gaese BH (2009) Correlating stimulus-specific adaptation of cortical neurons and local field potentials in the awake rat. J Neurosci 29:13837–13849. doi:10.1523/jneurosci.3475-09.2009

Warren RM (2008) Auditory perception: an analysis and synthesis, Third edn. Cambridge University Press, Cambridge. doi:10.1017/CBO9780511754777

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. doi:10.1093/beheco/arv204

Winkler I, Denham S, Mill R, Bohm TM, Bendixen A (2012) Multistability in auditory stream segregation: a predictive coding view. Philos Trans R Soc Lond B Biol Sci 367:1001–1012. doi:10.1098/rstb.2011.0359

Zelick RD, Narins PM (1982) Analysis of acoustically evoked call suppression behavior in a neotropical treefrog. Anim Behav 30:728–733. doi:10.1016/S0003-3472(82)80144-9

Chapter 4
Mate Searching Animals as Model Systems for Understanding Perceptual Grouping

Hamilton E. Farris and Ryan C. Taylor

Abstract A critical component of communication in humans and nonhuman animals is the ability to group signals so that they can be assigned to their correct sources. This is especially true for mate choice behavior, as incorrect stimulus grouping could lead to inaccurate evaluation of signalers by receivers, ultimately resulting in costly mate choice decisions. Sexual signals are often complex, consisting of components that vary in several physical parameters and across sensory modalities. Thus, the mate choice behavior of receivers is well suited for psychophysical tests of the limits and mechanisms of perceptual grouping both within and across sensory modalities. This chapter examines perceptual grouping in comparative models of mate choice behavior. We focus primarily on mate attraction in frogs, reviewing first the effects of spectral, temporal, and spatial parameters on sequential and simultaneous auditory grouping. We then review research on cross-modal perceptual grouping of frog visual and acoustic signals, a perceptual ability analogous to that of grouping human speech with its coincident mouth movements. In addition, we suggest that data from comparative models are not only useful for understanding signal processing in animal communication but also for potentially understanding the fundamental mechanisms receivers use to sort complex signals across all taxa and how such mechanisms may evolve.

H.E. Farris (✉)
Neuroscience Center, Louisiana State University Health Sciences Center, New Orleans, LA, USA
e-mail: hfarri@lsuhsc.edu

R.C. Taylor
Department of Biology, Salisbury University, Salisbury, MD, USA
e-mail: rctaylor@salisbury.edu

M.A. Bee, C.T. Miller (eds.), *Psychological Mechanisms in Animal Communication*, Animal Signals and Communication 5,
DOI 10.1007/978-3-319-48690-1_4

4.1 Introduction

Sensory systems across all taxa are bombarded with a variety of stimuli conveying information about the environment, in general, and the sources or producers of those stimuli, in particular. To evaluate this information accurately, stimuli must be correctly grouped based on their sources, such that stimuli produced by the same source are perceived to be members of the same group (Chap. 3). Perceptual grouping is a requirement across a range of functional contexts, including communication. Because receivers process communication signals that evolved to convey specific information (Bradbury and Vehrencamp 1998), incorrect stimulus grouping or source assignment would lead to misinformation (Bregman 1990; Fay 2008), increasing the likelihood of incorrect and potentially costly receiver decisions. In the broad context of mate choice, such costs could arise due to hybrid matings resulting from failures to recognize the signals of one's own species, failures to choose a high-quality mate among the conspecific mates available, and failures to accurately assess the size and fighting ability of competitive rivals (Maynard Smith and Harper 2003). Given the potential evolutionary significance of grouping, we should generally expect receivers to possess perceptual and cognitive mechanisms for doing so accurately.

How the brain "sorts out" (Darwin 1997) which stimuli should be grouped is an old problem faced by ancestral sensory systems and those derived from them (Popper and Fay 1997; Fay and Popper 2000). Consequently, extant mechanisms for perceptual grouping across species potentially reflect both ancient mechanisms inherited from past common ancestors as well as novel mechanisms that have been more recently derived and are unique to specific lineages (Bee and Micheyl 2008). The broad perspective of this chapter is that integrating comparative psychophysical and neuroethological studies can reveal whether aspects of perception and cognition are taxonomically widespread as a result of sharing neural mechanisms inherited from a common ancestor or more recently evolved in a more limited number of species. This comparative approach can generate predictions about how underlying mechanisms for perceptual grouping may (or may not) differ across taxa based on the extent to which those taxa share perceptual traits and sensory ecology (Kurylo et al. 1997; Burke et al. 2001; Spinozzi et al. 2004, 2009; Neiworth et al. 2014). For example, given similarities in visual ecology and visual mechanisms across many vertebrates, we should not expect our own abilities to group visual stimuli perceived to be in close spatial proximity (Rock and Brosgole 1964; Rock and Palmer 1990) to be unique to humans. Integrating comparative psychophysical and neuroethological studies can help inform our understanding of mechanisms for perceptual grouping by also explicitly considering the ecological and functional contexts in which different taxa make grouping decisions (see also Chap. 3). This perspective is important because, from a comparative point of view, studies of perceptual grouping in humans assay perceptual and cognitive systems in a manner that is often removed from the contexts in which they most likely evolved. Whereas experimental designs in human psychophysics are often very good at measuring stimulus features that elicit grouping, they do not generate predictions as to why certain grouping capabilities have been favored over others. In contrast, research that is based on the functional context of perception can lead to experiments that address mismatches between experimentally measured perceptual capabilities and those expected to operate in nature (Farris and Ryan 2011). An additional advantage of using a comparative approach is that many nonhuman taxa can provide direct access to the neural substrates so that the neural mechanisms underlying grouping decisions may be assayed experimentally (Chap. 3).

In taking a comparative approach, it becomes necessary to ask, what are the criteria for measuring grouping in an animal model? Whether using learned or natural behavior in experiments, behavioral responses indicating grouping must be distinguishable from responses to the individual stimulus components when presented alone. For example, grouping may be operationally defined as follows: for the simplest of groups containing two stimuli, A and B, responses must differ for each stimulus when presented alone and in combination. Thus, stimulus A elicits response A, stimulus B elicits response B, and stimulus AB elicits response C. This can be illustrated by the perception of an auditory stream, a sequence of sounds perceived to be produced by the same acoustic event (Bregman 1990; Chap. 3). In a classic example of auditory streaming (see Sect. 4.2.2), two sequences of repeating tones differing in frequency are interleaved (ABABAB. . .) (Miller and Heise 1950; van Noorden 1977; Chap. 3). When the frequency difference is sufficiently large, listeners report hearing two distinct streams, one of just A tones and one of just B tones, that are similar to when sequences of A or B tones are presented alone. In contrast, when the frequency difference between the A and B tones is small, the percept is of a grouped, single stream that rapidly alternates back and forth between the two frequencies.

One behavior that is both functionally significant and experimentally well suited to investigate grouping is mate choice, in which receivers respond to stimuli with source-specific behavior, such as phonotaxis (Walker 1957; Gerhardt and Huber 2002). Both intraspecific and interspecific selection (e.g., resulting in reproductive character displacement) have favored the evolution of signals with limited or relatively distinct sets of stimulus components. For example, selection against heterospecific mating that produces low-fitness hybrid offspring favors the evolution of signals with acoustic properties, such as dominant frequency or temporal pattern, that are species specific (Walker 1974; Shaw and Herlihy 2000). Consequently, when receivers are presented with signals from many different species, they must perceptually segregate signals so that responses are directed to those containing species-specific components rather than a mixture (Gerhardt 1994). It is well known, of course, that such responses to sexual signals occur across disparate taxa that use a wide variety of signal components in different sensory modalities (Andersson 1994). Thus, not only is the phenomenon of perceptual grouping likely widespread, it is also readily performed during the execution of adaptive behaviors in which correct grouping and source assignment are critical: sexual signaling and mate choice. Furthermore, with respect to experimentation, perceptual grouping during mate choice behavior often does not require conditioning (e.g., go/no-go tasks; Chap. 3), as a species' inherent mating behaviors lend themselves well to repeated testing on stimulus perception.

Our focus in this chapter is on studies of perceptual grouping during mate choice primarily in anurans (frogs and toads) and, to a lesser extent, orthopterans (e.g., crickets). This focus allows us to build on previous reviews of scene analysis in nonhuman taxa, which predicted much of the grouping behavior in anurans and orthopterans reviewed here (Hulse 2002; Yost et al. 2008). To the extent possible, we review studies that are based on approaches from comparative psychophysics and neuroethology and that have sought to determine the stimuli and mechanisms that are necessary and sufficient for perceptual grouping in their adaptive context. The chapter is organized into two larger sections, one reviewing work on unimodal grouping of acoustic stimuli (Sect. 4.2) and a second reviewing studies of cross-modal grouping of acoustic and visual stimuli (Sect. 4.3). We end by summarizing the chapter and discussing directions for future studies (Sect. 4.4).

4.2 Auditory Grouping

Sound is a flux of mechanical energy, transferring movement from a source to particles of a medium. That fact that sound does not propagate in a vacuum (it requires a medium) is significant to the phenomenon of auditory grouping, as movement of particles in a medium is not limited to sounds from only one source or even one type of sound from a single source. Indeed, particle movement is not source specific, but rather is the sum of all sounds relative to time and distance. This means that auditory systems must deconstruct complex (summed) sound pressure waves into their component parts so that they can be grouped according to their sources. As noted in previous reviews focused on human auditory perception (Yost and Sheft 1993), this process of deconstruction and subsequent grouping has been variously termed auditory scene analysis, auditory image analysis, and auditory object perception (Bregman 1990; Hartmann and Johnson 1991; Moore 2012). Because this review will cover both unimodal and cross-modal grouping, we will more generally refer to this process as perceptual grouping.

Following the publication of Bregman's (1990) seminal book on auditory scene analysis, the concept of perceptual grouping and its role in perception of acoustic scenes has significantly influenced research on complex auditory processing in nonhuman models. For example, prior to Bregman's book, most auditory research on comparative animal models focused on the limits of parameter processing, such as detection and discrimination in the spectral (e.g., critical bands, tuning curves), temporal (e.g., amplitude modulation, integration), and amplitude (measuring audibility through absolute thresholds, masking) domains (Fay 1988). Indeed, auditory scene analysis and pattern perception based on perceptual grouping were largely considered only in the context of human experimentation (Dooling and Hulse 1989). Following Bregman's book, however, a larger segment of auditory research in nonhuman subjects has clearly demonstrated perceptual grouping that is often based on similar stimulus parameters or grouping cues to those used in humans (Bee and Micheyl 2008; Bee 2015; Chap. 3).

4.2.1 Auditory Grouping Cues

There are several physical stimulus parameters that can cue stimulus inclusion or exclusion in perceptual groups. Based on Gestalt principles, Bregman (1990) described two categories of grouping cues: primitive and schema based. For this review, we will define primitive cues as those that may be available in any sound, including arbitrary or nonfunctional ones, presented across a variety of stimulus contexts. They include spectral cues, such as common fundamental frequency, harmonic relationships (harmonicity), and bandwidth; temporal and amplitude cues such as common onset and amplitude modulation; and spatial cues such as common interaural level and time (phase) differences (van Noorden 1977; Darwin and Carlyon 1995; Darwin 1997; Moore and Gockel 2002; Hawley et al. 2004; Elhilali et al. 2009). Primitive grouping is often thought to be based on innate and potentially pre-attentive processing of such acoustic data. In contrast, schema-based grouping is often considered to employ learned stimulus associations and transition probabilities, such as in speech, and may include endogenous, attention-driven biases (Bregman and Rudnicky 1975; Bronkhorst 2015). But these Gestalt-based criteria for primitive and schema-based grouping are largely based on work with humans, leading to some question as to whether (or to what extent) schema-based grouping may be innate versus learned (Bregman 1990). The literature reviewed in this chapter demonstrates that when answers to this question also consider data from nonhuman animals, it is clear that innate, schema-based mechanisms may also mediate perceptual grouping of the nonarbitrary, complex sounds used in animal communication. In other words, grouping that is not driven solely by processing similarities in acoustic data can still result from innate mechanisms, which are potentially mediated by innate perceptual templates for complex groups (Capranica and Moffat 1983). We review such a case in Sect. 4.2.2.1 and later discuss neurophysiological data that indicate low-level processing of primitive cues by the ascending auditory system is under descending control, providing evidence to suggest primitive cues are not exclusively processed pre-attentively (Farris and Ryan 2011; Ponnath and Farris 2014).

4.2.2 Auditory Streaming

Sequential auditory grouping refers to assigning sounds produced at different times to the same source. It is a critical form of stimulus segregation due to the nature of sound production in animals. Whether produced by percussive or respiratory mechanisms, sequences of sound result from the discontinuous but repetitive driving of sound producing structures. A sequential signal results when the behavior (i.e., the muscle pattern) resets to restore the driving force, such as refilling air (Gans 1973) or repositioning a percussive tool (see Gerhardt and Huber 2002). This creates amplitude modulation, or times with sound on and sound off. Sequential grouping necessitates an auditory memory task, as the ongoing sensory responses to the components of sound mixtures are compared to those elicited in some previous time window. Those components perceived as being most similar in some property elicit the perception of a grouped sequence or an auditory stream.

The perception of auditory streams is perhaps the most easily demonstrated form of sequential grouping. As introduced in Sect. 4.1, when presented with a sequence of tone pulses of alternating high and low frequencies (e.g., ABABAB. . .), human listeners perceive two separate sequential groups of constant frequency provided the frequency difference is sufficiently large (e.g., frequency ratios > 1.15): a stream of high-frequency pulses (e.g., A–A–A–. . .) and a stream of low-frequency pulses (e.g., –B–B–B. . .). In contrast, smaller frequency differences (e.g., frequency ratios < 1.15) elicit perception of a single stream of pulses occurring at twice the pulse rate with its frequency jumping up and down (Miller and Heise 1950). The frequency ratio required for perceptual fission/fusion of the two interleaved sequences is related to the frequency resolution of the auditory periphery (Rose and Moore 2000). This relationship suggests that low-level auditory filtering, or the segregation of stimulus spectral components into separate frequency channels, underlies the grouping percept (Fishman et al. 2001, 2004; Bee and Klump 2004, 2005; Pressnitzer et al. 2008; Chap. 3). In addition to frequency differences, other primitive cues such as temporal structure (e.g., pulse rate, duration, interpulse interval), amplitude similarity, and spatial separation have been shown to affect the streaming percept (Bregman and Campbell 1971; van Noorden 1977; Bregman 1990; Moore and Gockel 2002, 2012; Cusack and Carlyon 2004). When compared to the acoustic structure of human speech, tone streaming based on frequency similarity and pulse rate is considered a relatively simple form of auditory processing (of course, processing of many different primitive cues may be employed together during sequential grouping of acoustically complex sounds such as speech; Darwin and Carlyon 1995). Indeed, for organisms using acoustically complex signals (e.g., speech and birdsong) such pure tone stimuli on their own may be ecologically rare, making the performance of tone streaming a laboratory artifact created by a linear systems approach to experimental design (Bregman 1990; but see Chap. 3). But from a comparative point of view, many animal signals are quite similar to a series of constant frequency pulses. Furthermore, they are produced amidst overlapping signals consisting of tone pulses of different frequencies and pulse rates (Nityananda and Bee 2011). Thus, exploring the use of primitive cues during streaming (as in tone streaming) in other taxa has the potential to elucidate processing that is fundamental to all auditory systems, especially when comparing grouping behavior that evolved independently.

4.2.2.1 Spatial Cues for Streaming

In many anuran species, males produce calls that function as sexual advertisement signals and that females and other males use to adjust mating behavior. In most taxa, mate choice is mediated by female phonotaxis: the movement of females to a call's source. Phonotaxis requires recognition and localization of the calls (Gerhardt and Bee 2007). Although the acoustic structure of calls shows extensive interspecific variation, including broadband and harmonically rich components produced with different temporal structures (Ryan 1985; Larson 2004; Feng et al. 2006), many species produce calls consisting of sequences of narrow-band (often tonal) pulses at consistent pulse rates (Martin 1971; Forester and Czarnowsky 1985; Gerhardt 2001). Given this call structure, receivers in a breeding aggregation, where the auditory scene consists of many males producing concurrent calls, face an analogous grouping task to that of human subjects in the classic streaming paradigm (Nityananda and Bee 2011). For example, imagine a simple scenario in which a female frog finds herself in a chorus of two males producing pulsatile calls. Consider a situation in which the two males produce their calls with extensive but imperfect overlap, such that the pulses of one male's calls were interdigitated with the pulses of the other male's calls. (This can actually happen in frogs.) Because pulse rate is often an important species recognition cue, correct segregation of the two calls into two separate streams should elicit phonotaxis to one of the two callers, whereas fusion of the two calls into one stream should not elicit phonotaxis. Interpretation of the latter response is that there is perception of a single caller producing pulses at twice the species-specific pulse rate and thus, an inappropriate call, possibly from another species. Schwartz and Gerhardt (1995) used this exact experimental setup to test the spatial limits of stream segregation in the eastern gray treefrog, *Hyla versicolor*. They found that although increases in spatial separation between two pulsatile calls increased the likelihood of phonotaxis, females still exhibited evidence for grouping over large spatial separations (120°). In a clever switch to this design, Bee and Riemersma (2008) tested the spatial limitations to sequential grouping in Cope's gray treefrog, *Hyla chrysoscelis*. Rather than investigating the separation at which two full calls segregate, they measured the spatial separation required for two interleaved calls at half the species-specific pulse rate to fuse. Again, although there was more evidence for grouping at smaller spatial separations, frogs still showed grouping up to 180°. Considering the behavior in its functional context (i.e., making a mate choice decision), such poor spatial resolution in grouping would seem paradoxical (i.e., poor spatial resolution would correlate with poor mate discrimination). We will return to this question below in some of our own work, which may resolve, in part, this paradox.

The use of this split call approach is not limited to tests with vertebrates. In most species of true crickets (Family: Gryllidae), a similar pulsed call is produced when males stridulate by engaging specialized structures on each of the forewings as they are closed across each other (Bennet-Clark 1989). A single closure engages the file and scrapper mechanism that delivers a driving frequency to coupled resonant parts of the wings. The result of a single closure is a short duration (~7 – 12 ms) tonal pulse with Q_{3dB} values of 10–20 [note that other orthopterans, such as tettigoniids, use similar mechanisms, but often produce more broadband sounds (Bennet-Clark 1995, 1999; Dambach and Gras 1995; Gu et al. 2012)]. Both the temporal patterns of wing closure and the resonant (i.e., carrier) frequency vary strongly between species and between individuals (Walker 1962). Across cricket species, females exhibit strong preferences for specific temporal patterns and carrier frequencies, meaning phonotaxis is only elicited if subsequent pulses in a series are perceived to have the appropriate timing and frequency (Walker 1957; Pollack et al. 1984; Schildberger 1984; Shaw and Herlihy 2000). Such a preference can only occur if the sequential pulses produced by a male are perceptually grouped as part of the same stream. Thus, positive phonotaxis to one source among many indicates grouping, enabling another ecologically valid application of the classic streaming test. For example, although rarely discussed in terms of streaming, work by Weber and Thorson (1988) used phonotaxis in field crickets (*Gryllus bimaculatus*) to measure grouping of two sequences of pulses that were experimentally interleaved and spatially separated by 0–180°. Depending on their temporal structure, the stimuli included variations in which the species-specific call characteristics were produced by individual speakers or the composite of the two (e.g., temporally interleaving two songs, each having half the pulse rate of male songs). As with both species of gray treefrogs, when the composite temporal parameters of the spatially separated calls matched the species-specific pattern, females performed phonotaxis, providing evidence of call recognition, and, thus, grouping. Clear evidence for grouping persisted up to angular separations of 135°, again similar to results from gray treefrogs.

Considering the data above from a functional perspective raises a problem. The large grouping angles measured would seem to create the possibility for mate choice errors in a chorus. Large grouping angles would create little confidence in the receiver as to which pulses belong to the same call and thus a particular male. Farris and Ryan (2011) addressed this question in túngara frogs (*Physalaemus pustulosus*) by testing the hypothesis that relatively permissive auditory grouping is an artifact of presenting only a limited number of stimuli: the fewer the available stimuli, the more likely dissimilar ones will be grouped. As the number of stimuli increased, group membership was predicted to become more selective. A short description of the stimuli used in these tests with túngara frogs is necessary, as they are more acoustically complex than those used in tone or narrow-band streaming. Túngara frogs differ from the taxa reviewed above in that males, rather than producing a sequence of repeated narrow-band pulses in their calls, produce a call consisting of a 350-ms, frequency-modulated (FM) sweep (dominant frequency: ~900–400 Hz) called the "whine" followed by 0–7 harmonic bursts (40–80 ms duration) called "chucks" (Fig. 4.1a; Ryan 1985). The acoustic complexity of these components approaches that known to be streamed by birds (Braaten and Hulse 1993; Hulse et al. 1997; MacDougall-Shackleton et al. 1998) and potentially mice (Gaub and Ehret 2005). And, as the onomatopoeic name suggests (túngara: /toon/·/guh/·/ruh/), they are arguably speechlike in acoustic

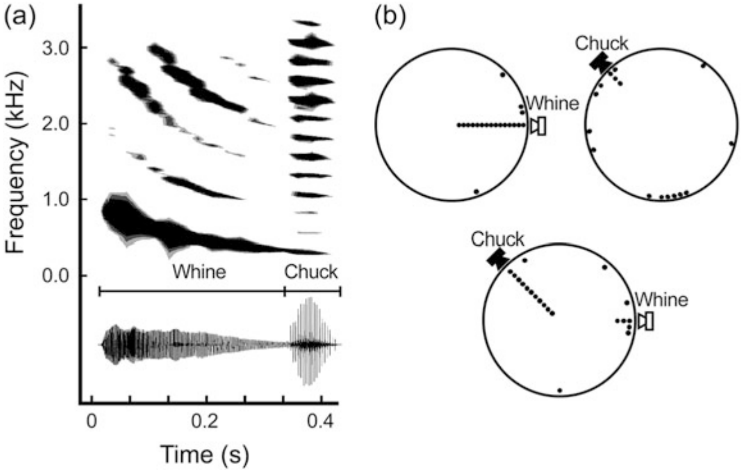


Fig. 4.1 Vocalizations and perceptual grouping in the túngara frog (*Physalaemus pustulosus*). (**a**) Spectrogram (*top*) and waveform (*bottom*) of the túngara frog's advertisement call illustrating a whine following by a single chuck. (**b**) Behavioral results from a phonotaxis experiment showing the permissiveness of perceptual grouping in túngara frogs when receivers are not given opportunities to make relative comparisons. Points depict exit angles from a circular arena in response to a whine alone, a chuck alone, and a whine and chuck having the natural temporal sequence but separated by 135° (Farris et al. 2002).

structure, resembling vowel-like and consonant-like components, respectively. Furthermore, there is a rudimentary form of syntax, in that the whine-chuck sequence is constrained by laryngeal morphology (Ryan and Drewes 1990; Gridi-Papp et al. 2006), meaning chucks can only follow whines. Whereas the whine is necessary and sufficient to elicit and direct phonotaxis, the chuck alone (i.e., artificially presented alone) does not elicit phonotaxis. However, the chuck does affect preference, as the addition of the chuck to a whine increases the whine's attractiveness relative to a whine alone (Ryan and Rand 1990). For a female to implement this preference in a multimale chorus, she must accurately determine which whine goes with which chuck, an acoustically analogous problem to that of the cocktail party problem (Cherry 1953; Cherry and Weary 1954; Bee and Micheyl 2008).

To measure whine-chuck grouping, Farris et al. (2002, 2005) established an assay that satisfies the grouping criteria outlined earlier in Sect. 4.1: a whine presented alone elicits phonotaxis (response A); a chuck alone elicits no phonotaxis or random movement (response B); and a whine presented with a spatially displaced chuck elicits phonotaxis to the chuck (response C). This response to the chuck, which is conditional to the presence of a whine, can then be used to measure the spatial, temporal, and amplitude criteria for whine-chuck grouping. When presented with a single whine and chuck in the natural temporal sequence at the natural rate (whine-chuck at 2-s period), grouping was measured up to 135°

whine-chuck separations, similar to those measured in studies of gray treefrogs and field crickets (Fig. 4.1b; Farris et al. 2002). If we extrapolate these data to the field, we would predict wide grouping angles and thus poor grouping accuracy. However, when presented with a single whine followed by the same chuck broadcast from two different speaker locations, grouping is more likely for the whine-chuck with the smallest relative spatial separation (Farris and Ryan 2011). That is, even though either chuck could have been grouped, group membership was based on a relative comparison of spatial separation. It is not yet known whether this relative spatial comparison was based on interaural level or interaural time differences (or both). It is interesting to note that this increase in spatial acuity during grouping differs from the spatial acuity of noise masking in túngara frogs, a result consistent with those in humans showing differences in the ability to spatially segregate speech as compared to noise (Freyman et al. 1999, 2001; Farris and Ryan 2011). This difference, when viewed from a comparative perspective, suggests that specialized segregation of complex communication sounds may not be limited to human speech, but rather may be more associated with communication sounds in general, especially those vocalizations produced in multisource environments.

Consistent with the hypothesis that grouping of communication signals employs perceptual mechanisms specialized for the species-specific stimuli is the fact that relative comparisons during grouping in túngara frogs are not limited to a primitive cue (spatial). Grouping by túngara frogs also employs schema-based comparisons. As noted earlier, a whine precedes a chuck if both are produced by the same, single individual. Thus, in a chorus, at least one schema-based grouping cue for females is based on the whine-chuck temporal sequence: valid groups should only contain whines and chucks in the species-specific order. If, for example, there is whine-chuck temporal overlap, they cannot be produced by a single male and thus should not be grouped. There is no acoustic rationale for this grouping cue other than the syntactic schema of the species-specific male call. As when a single whine and chuck are presented in various spatial orientations, a single whine and chuck presented with various temporal arrangements will elicit grouping. However, when presented with a single whine and two chucks with different relative timing to the whine, whine-chuck grouping is more likely for the chuck timed closest to the natural sequence (Farris and Ryan 2011). Taken together, the data suggest that even in relatively simple auditory systems (e.g., those lacking cortical circuitry), there exists the capability for the analysis of common primitive cues along with the analysis of complex signal structure, including transition probabilities of acoustically complex signals.

4.2.2.2 Spectral Cues for Streaming

Given that mate choice involving phonotaxis inherently has a directional component to the behavior, tests of nonspatial grouping cues still use a spatially separated experimental design to distinguish responses. Thus, Nityananda and Bee (2011) exploited the split call experimental design to test whether the frequency of interleaved pulses affected phonotaxis in Cope's gray treefrogs (*Hyla chrysoscelis*). A pulsatile target call, which is an attractive stimulus, and a pulsatile distractor were presented from the same position around a circular arena such that their pulses were temporally interleaved. When the carrier frequency of the distractor was similar to that of the attractive target, phonotaxis to the target was reduced, suggesting poor segregation of the two stimuli, such that they were processed as an integrated stream with temporal parameters presumably twice the rate of natural calls. As the difference between the carrier frequencies of the two stimulus alternatives increased, however, females showed greater attraction to the target, a result analogous to the fusion boundaries measured in the ABAB streaming paradigms (Moore and Gockel 2002), including that for the informational masking of speech in human listeners (Darwin et al. 2003).

The acoustic structure of the túngara frog's whine and chuck allows for testing spectral influences on grouping for more complex stimuli. In particular, given the broadband spectrum of the two call components (Fig. 4.1a), experiments can ask to what extent does peripheral channeling play a role in the conditional response to the chuck. Although filtering stimuli to enhance particular frequency bands is quite easy in such experiments, certainty that potential grouping components are filtered to different peripheral frequency channels may be difficult in nonhuman taxa, in which careful measurement of peripheral frequency resolution has not been done. Frogs, however, facilitate such tests due to their auditory anatomy: there are two primary auditory end organs in the anuran inner ear, the amphibian papilla (AP) and the basilar papilla (BP), that process low-frequency and high-frequency sounds, respectively (Simmons et al. 2007). Thus, filtering call components to the sensitivity ranges of the two organs facilitates separate peripheral channeling of the stimuli in a manner arguably similar to dichotic stimulation (i.e., presenting different sounds to different ears). In túngara frogs, a neural audiogram based on multiunit neural recordings from the auditory midbrain suggest the AP is sensitive to frequencies less than 1.2 kHz, with BP sensitivity centered at about 2.1 kHz (Ryan et al. 1990). Farris et al. (2005) showed that whine-chuck grouping, although weaker than for full-spectrum stimuli, still occurs when the whine is low-pass filtered, leaving just its dominant frequency so that it stimulates only the AP, and the chuck is high-pass filtered to stimulate only the BP. Thus, analogous to dichotic grouping, the grouping response still occurs when sounds largely stimulate separate end organs and thus separate peripheral channels (i.e., in this case, spectral filtering reduces overlap of stimulation to the sensory epithelia). This means that central mechanisms appear sufficient for grouping and that no peripheral overlap or channeling (Hartmann and Johnson 1991) is necessary. Of course, across-frequency or across-channel grouping is not limited to sequential grouping and is commonly found during the perceptual grouping of stimuli presented simultaneously, which we address in the next section.

4.2.3 Simultaneous Grouping

Simultaneous grouping, in which concurrent stimuli with different physical parameters (e.g., frequencies) are assigned to the same perceptual group, is clearly relevant when considering that communication signals in many taxa are acoustically complex, requiring the auditory system to group *different* frequencies presented at *common* times (Bee and Micheyl 2008; Elhilali et al. 2009; Micheyl et al. 2010). From a mechanistic point of view, this means the auditory system must assess commonalities in ongoing neural activity produced from two different places on a single sensory epithelium (e.g., the basilar membrane in mammals) or produced from different sensory epithelia altogether (e.g., the AP and BP in anurans). Although we review only one example of simultaneous grouping across different auditory frequency channels, the influence of temporal coherence on grouping extends beyond the auditory system and includes cross-modal grouping of stimuli produced in different sensory modalities, which is the focus of Sect. 4.3.

Male Cope's gray treefrogs (*Hyla chrysoscelis*) produce calls with two frequency components near 1.1 and 2.2 kHz. These two components almost certainly stimulate the AP and BP, respectively (Hillery 1984; Schrode et al. 2014). Females prefer the two-frequency call over calls containing either single frequency alone (Bee 2010). Using this preference, Bee (2010) measured the spatial limitations for the simultaneous grouping of the two frequency components: females were given a choice between a complete call presented from a single speaker versus each frequency component presented from two spatially separated speakers. Phonotaxis was more likely to the call with the spatially coherent spectral components (full spectrum) than those components spatially separated into two single-frequency calls. Furthermore, this preference matched that found when females were given choices between the full-spectrum call and a single call with only part of the spectrum. Thus, the results suggest that females did not perform simultaneous grouping of the spatially separated call components, treating them instead as individual calls with only part of the spectrum. In contrast to some of the studies of sequential grouping discussed in Sect. 4.2.2, even small angles of separation (e.g., 7.5°) between simultaneous components shifted female preferences toward the call with two spatially coherent spectral components. Preliminary data reported in Bee (2015) from a study of green treefrogs, *Hyla cinerea*, suggests frogs also exploit common onsets/offsets as cues for grouping simultaneous signal components that separately stimulate the AP and BP.

4.3 Cross-Modal Grouping

In addition to grouping stimuli within a single sensory system, nervous systems also group stimuli by integrating information across sensory systems. Multisensory integration has been long appreciated in human perception and communication.

For example, sonogenic synesthesia, where music can generate distinct visual experiences in some people, has been known for over 100 years (Henson 1977). Additionally, almost everyone is familiar with the ventriloquism effect (Howard and Templeton 1966), where the source of speech appears to emanate from a dummy rather than the entertainer. Likewise, in human speech, the degree of coherence between phonemes and patterns of lip movement can influence auditory perception (McGurk and Macdonald 1976). An impressive body of work in psychophysics and neurobiology has revealed an array of mechanisms by which humans conduct multisensory integration (Stein 2012). Outside of human psychophysics, a substantial body of work on multisensory perception has been built through studies of cats (*Felis catus*), rhesus macaques (*Macaca mulatta*), and barn owls (*Tyto alba*) (Knudsen 1982; Knudsen and Knudsen 1989; Stein and Meredith 1993; Whitchurch and Takahashi 2006).

The past 15 years has also seen an explosion of research on multimodal (multisensory) signaling in a diverse array of nonhuman animals, including spiders (Elias et al. 2006; Hebets 2008; Kozak and Uetz 2016), fish (Maruska et al. 2012), anurans (Starnberger et al. 2013; Reichert et al. 2016), birds (Uy and Safran 2013), and mammals (Partan et al. 2009; Stein 2012). What most of these recent studies have in common is the examination of multimodal signaling through the lens of evolutionary biology, typically with respect to how female mate choice drives the evolution of complex signals (see Chaps. 5 and 6; see Munoz and Blumstein (2012) for a review of multisensory integration in predator-prey studies). Although important perceptual and cognitive information can be gleaned from these studies, like early auditory work, most were not designed to investigate directly how receivers group signal components across stimulus modalities (cross-modal grouping). Thus, from a comparative standpoint, we know very little about how nonhuman animals integrate their senses to form cross-modal perceptual groups. This limits not only our understanding of the diversity of multisensory integration but can also limit our understanding of how mammalian nervous systems, like our own, have evolved to generate perception.

One reason few researchers have addressed questions of multisensory perception in nonhuman animals may stem from the difficulty of interpreting how behavior reflects a receiver's perception. Fortunately, as with auditory grouping, some behavioral systems are tractable for addressing questions of cross-modal grouping; anuran amphibians (Bee 2015) and spiders (Kozak and Uetz 2016) are obvious choices. Both taxonomic groups are easily manipulated in experimental conditions and respond readily to multimodal playbacks. An impressive body of multisensory research has been done in spiders, but most of this work has focused on signal evolution, rather than addressing perceptual mechanisms and their evolution directly (but see Kozak and Uetz 2016). The breadth of knowledge that is now available in spiders (e.g., Herberstein et al. 2014) makes this group a prime candidate for future studies in perception. In frogs, some multisensory perceptual questions have been addressed directly. The remainder of this chapter focuses on work in this taxonomic group.

4.3.1 Anuran Multisensory Communication

The use of visual signals in anuran communication has been postulated for 35 years or more, especially with regard to color patterns as signals (Wells 1980; Sztatecsny et al. 2012). A number of studies have also examined specialized limb-waving (semaphoring) behaviors in frogs as an evolved visual communication system (Lindquist and Hetherington 1998; Amézquita and Hödl 2004; Preininger et al. 2013a, b). Compared with acoustic studies, however, we know very little about the role of vision in anuran communication. This has likely occurred because of the success of acoustic research and an assumption that visual information would be limited for a group of animals in which many species communicate primarily at night (Buchanan 1993). Many frog species, however, possess large, highly sensitive eyes that provide visual information even under very low-light conditions (Aho et al. 1988; Buchanan 1993; Cummings et al. 2008). In the early 2000s, researchers began to give serious consideration to the idea that visual communication may be important, even in anurans that lack specialized semaphoring behavior, particularly given that the vocal sac inflation of calling males may have been co-opted as an important communication cue (Narins et al. 2003; Taylor et al. 2007). Within the past 15 years, studies have revealed that a diverse array of anuran species use the vocal sac as a visual cue (Rosenthal et al. 2004; Hirschmann and Hödl 2006; Taylor et al. 2007, 2008; Preininger et al. 2013b; Starnberger et al. 2014a, b), and, in some species, to improve source localization (Narins et al. 2005). In most species, the acoustic signal seems to have primacy, but the visual component can strongly modulate female responses to acoustic signals. Despite the recent increase in available data on anuran multisensory communication, we still only have a rudimentary understanding of the processes by which these animals integrate multisensory information and how this integration influences perception, behavior, and signal evolution.

4.3.1.1 Cross-Modal Grouping in Túngara Frogs

Unlike many songbirds, frogs usually have a more limited repertoire of call types, making their relatively simple acoustic signals experimentally useful for quantifying the limits of perception. Here again we focus on the Neotropical túngara frog and its somewhat unusual two-note advertisement call (Fig. 4.1a). Employing a robotic frog that has an inflatable vocal sac (Fig. 4.2), Taylor et al. (2008) demonstrated that, all else being equal, female túngara frogs preferentially approach a call when they can also see a vocal sac inflating synchronously with a call emanating from a speaker behind the robotic frog (see Fig. 4.3d). That is, females preferred a bimodal stimulus (visual + auditory) over a unimodal (auditory) one. Similar to the chuck, the vocal sac, when presented in the absence of a whine, also fails to elicit signal recognition by females. Thus, for both chuck and the vocal sac responses, the whine must be present in order for females to recognize them as appropriate mating

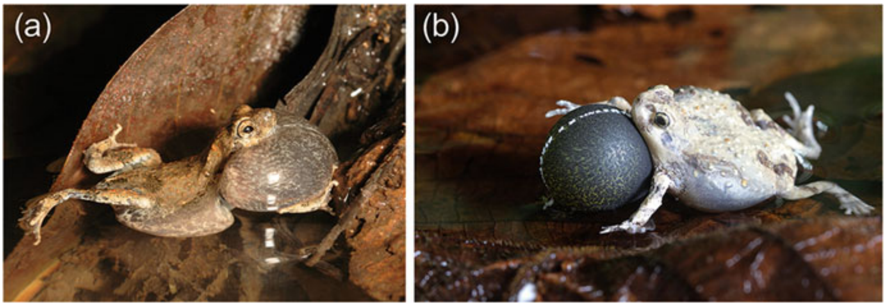


Fig. 4.2 Real and robotic túngara frogs (*Physalaemus pustulosus*). A real calling male túngara frog is depicted on the *left,* and a robotic túngara frog with inflatable vocal sac is depicted on the *right*

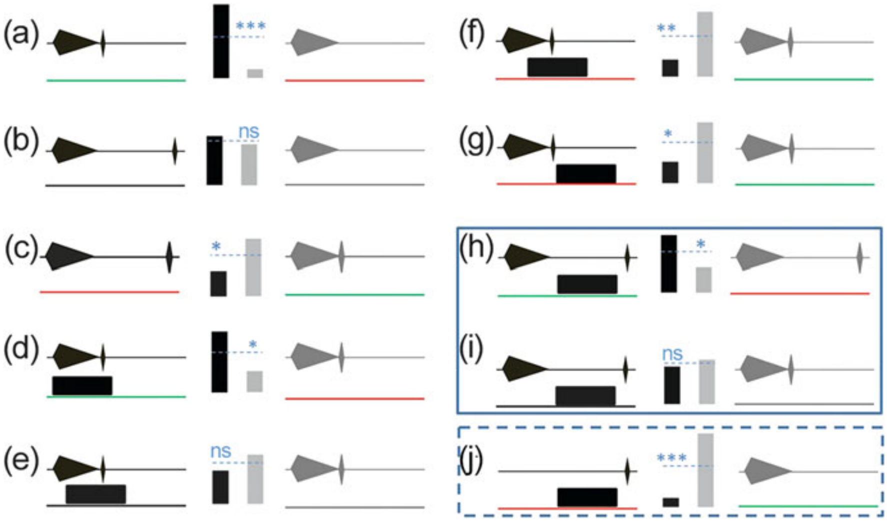


Fig. 4.3 Results from two-alternative choice tests comparing phonotaxis responses to various unimodal and bimodal stimuli in túngara frogs (*Physalaemus pustulosus*). In each of 10 tests (**a**–**j**), 20 females were given a choice between the signal in *black* (*left*) versus the signal in *gray* (*right*). The *vertical black* and *gray bars* for each test represent the number of females that chose the respective signal, and the *blue dashed horizontal lines* represent the null hypothesis of equal preference. Experiments highlighted in the *solid blue box* are tests of the perceptual rescue hypothesis, and those in the *dashed blue box* are the test of the component substitution hypothesis. Each portion of the figure schematically illustrates the acoustic components of the túngara frog mating call offered in each alternative: a whine only (**a**, **b**, and **j**; *right*), a chuck only (**j**, *left*), or a whine-chuck (all other alternatives). Tests in which one alternative was the whine-chuck in a natural sequence are depicted in (**a**) (*left*), (**c**) (*left*), (**d**–**g**) (*both alternatives*), and (**i**) (*right*). *Horizontal black rectangles* represent the inflation-deflation cycle of the robotic frog's vocal sac and its temporal relationship to the call in (**d**–**j**) (*black, left*). The *x*-axis depicts 1 s, with *green* indicating a significantly preferred stimulus and *red* indicating the less preferred stimulus. The results of binomial tests are noted as $^{***}P < 0.001$, $^{**}P < 0.01$, $^{*}P < 0.05$, and ns (not significant, $P > 0.05$)

signals. In a second set of experiments, Taylor et al. (2011) demonstrated the importance of temporal synchrony between vocal sac movement and the call being broadcast from the speaker. In this study, the investigators altered the timing of the vocal sac inflation such that it lagged the onset of the call by various degrees, something that males are physically unable to do in nature. When the vocal sac inflation lagged the call onset by 100 ms or less (e.g., the call and vocal sac inflation exhibited at least 75 % overlap), there was no significant decrease in female response to the bimodal signal. When the vocal sac inflation lagged the call onset by 200 ms or more (50 % overlap or 0 % overlap), females reversed their preference and preferentially approached the speaker with no visual stimulus associated with it (see Figs. 4.3f, g). In a follow up study, Taylor et al. (unpublished data) presented females with asynchronous stimuli in which the call lagged the visual stimulus. Females failed to show a significant aversion to the asynchronous bimodal stimulus in this case. Interestingly, this behavior in túngara frogs closely parallels human audiovisual integration. Dixon and Spitz (1980) showed that when viewers observed decoupled audiovisual film tracks, lip movements (visual cues) needed to lag speech onset by approximately 250 ms in order for asynchrony to be perceived. They also showed that when auditory cues lagged visual cues, this reduced perceptions of asynchrony compared to the reverse, a finding confirmed again by Bushara et al. (2001). Finally, Dixon and Spitz (1980) found that viewers were better able to detect asynchronies when they viewed and heard a hammer hitting a peg than for speech. Thus, incongruence in simple audiovisual stimuli was better resolved than in more complex stimuli such as speech.

These studies demonstrated that even in an acoustic specialist (a frog), visual cues can modulate receiver responses to the acoustic signal component. There are at least two potential explanations for these results. First, the vocal sac as a visual cue could be a sexual ornament, such that the mere addition of another component increases the overall attractiveness of the signal. Deviations in synchrony between audiovisual components could elicit reduced behavioral responses because the complex signal fails to match the animal's rigid internal template for appropriate species-specific signal parameters. From an organismal and evolutionary perspective, this view is sufficient to explain how complex signals evolve, but it glosses over important questions about sensory mechanisms, answers to which are critical for a full understanding of why signals evolve (Chaps. 5 and 6) and how nervous systems solve problems related to perceptual grouping (Chap. 3) .

The responses of female túngara frogs to incongruent bimodal signals also hint that more complex perceptual issues are at play. Females will not approach a robotic frog with inflating vocal sac when no acoustic signal is present, but when the vocal sac inflation follows the call with no temporal overlap, females recognize it as an incorrect signal and avoid it. Thus, the frogs exhibit sequential cross-modal grouping for at least a short period of time after the offset of the call. To investigate this further, Taylor and Ryan (2013) combined a series of audiovisual components in unnatural ways. First, they created a silent gap of 500 ms between the whine and chuck. They presented these separated call components to females from one speaker and played a whine only from the alternative speaker. Females responded equally to both speakers, indicating that they could no longer recognize that the temporally displaced chuck belonged to the whine; they effectively treated both signals as a whine (Fig. 4.3b). Because females only approach a chuck in the presence of a whine, this demonstrated that 500 ms is sufficient to eliminate sequential grouping in the acoustic domain, at least with respect to recognizing individual acoustic signal components. The next experiment placed the robotic frog's vocal sac inflation in the 500 ms gap between the whine and chuck so that the temporal order was whine, vocal sac inflation, then chuck. The alternative speaker broadcast the same separated call without a robotic frog. In this experiment, females approached the bimodal signal significantly more often than random, suggesting that the visual component perceptually rescued (i.e., formed a perceptual object of) the temporally separated whine-chuck acoustic components (Fig. 4.3h). An alternative explanation is that in the face of two unnatural signals that females do not experience in nature, the visual signal component merely elicited a favorable response when the alternative signal lacked such a feature. To test this, the authors then compared responses to the temporally disjunct bimodal signal to the normal, highly attractive whine-chuck. Here, the females responded equally to both signals, providing further indication that the "sandwiched" visual component generated a perceptual grouping of the temporally separated acoustic components (Fig. 4.3i). Another alternate explanation for this phenomenon is component substitution. Under this hypothesis, it is not that the visual cue generates perceptual binding, but rather that when the visual cue precedes the chuck, it effectively replaces the whine and renders the whine unnecessary. To test this possibility, Taylor and Ryan (2013) offered females a choice between an inflating robotic frog that immediately preceded the chuck and a speaker broadcasting just the whine. If the visual cue replaced the whine, then females should have preferentially approached the robotic frog and chuck combination. They did not. All but one female approached the isolated whine, indicating that the robotic frog and chuck combination did not elicit recognition of an appropriate mating signal (Fig. 4.3j). These data in túngara frogs indicate that the animals conduct sequential cross-modal grouping. In particular, receivers are attentive to signal components for a short period of time following the offset of the whine. The perception and subsequent response to the whine can then be disrupted or restored depending on both the type and sequence of signal components that follow.

Túngara frogs do not seem to group multisensory signal components when the acoustic and visual components are spatially separated. When a robotic frog was displaced from a speaker by a distance of only 15 cm, female receivers responded strongly to the speaker, almost completely ignoring the visual stimulus (Taylor et al. 2008). This result may not seem surprising in a group of animals where acoustic signals have primacy. However, another Neotropical frog species, the brilliant-thighed poison frog (*Allobates femoralis*), does group bimodal signals with spatially separated acoustic and visual components (Narins et al. 2005). In this diurnal species, males maintain territories and actively vocalize and defend territories against intruders. In that study, Narins and colleagues introduced a robotic male frog and speakers broadcasting vocalizations into the territories of resident male frogs. A 12-cm separation between the visual stimulus of the robotic male and the speaker broadcasting the call resulted in strong attacks toward the model instead of the loudspeaker. At greater distances of 25–50 cm, males began to direct their attacks toward the loudspeaker, but even a 50-cm separation was insufficient to fully eliminate spatial grouping. In a bird, the pied currawong (*Strepera graculina*), receivers also conduct multisensory spatial grouping. Lombardo et al. (2008) showed that the distance between the loudspeaker and bird model influenced receiver responses in this territorial bird. When the speaker and model were close together, currawongs approached the stimulus more closely than to either speaker alone or to the distantly separated speaker and model. The data from these three animal systems demonstrate species differences in strategies for cross-modal grouping. Perhaps it is not surprising that a frog and bird integrate streams of audiovisual information differently. What is interesting is that two species within the same order (Anura) seem to have evolved different strategies for cross-modal grouping. These differences may have evolved in response to the different ecological niches the species inhabit. In the túngara frog, females often have to discriminate among multiple male frogs that are calling within close proximity; 50-cm spacing among males is not unusual. Thus, if females were permissive about multisensory spatial grouping, effectively accepting a ventriloquism effect, they could more easily assign a call to the wrong source. On the other hand, males of *A. femoralis* tend to maintain spacing over much larger distances (many meters); thus, audiovisual spatial discrepancies are less likely to result in the misassignment of a call to its source. By taking a functional and comparative approach, investigators can measure and potentially explain differences in complex stimulus processing.

4.3.1.2 Cross-Modal Grouping in Treefrogs

Data from two North America treefrog species may also reflect divergent perceptual processes regarding cross-modal grouping. Cope's gray treefrogs (*Hyla chrysoscelis*) employ a variety of strategies in auditory scene analysis, including dip listening, and spatial and frequency segregation (Bee 2010; Nityananda and Bee 2011, 2012; Vélez and Bee 2011). For dip listening, female gray treefrogs had relatively lower signal recognition thresholds in the presence of slowly fluctuating sinusoidal amplitude-modulated (SAM) background noise compared with unmodulated noise. That is, when background masking noise fluctuated in amplitude at rates of less than 5 Hz, females experienced a 2–4 dB release from masking. Females effectively took advantage of the temporary decrease in masking noise to catch "acoustic glimpses" of individual male signals. Interestingly, the green treefrog, *Hyla cinerea*, failed to exhibit such an effect, suggesting that this species does not employ dip listening as a strategy for improving signal recognition (Vélez et al. 2012).

One possible explanation for the difference may be the process by which each species integrates visual and acoustic cues. Although the two species are closely

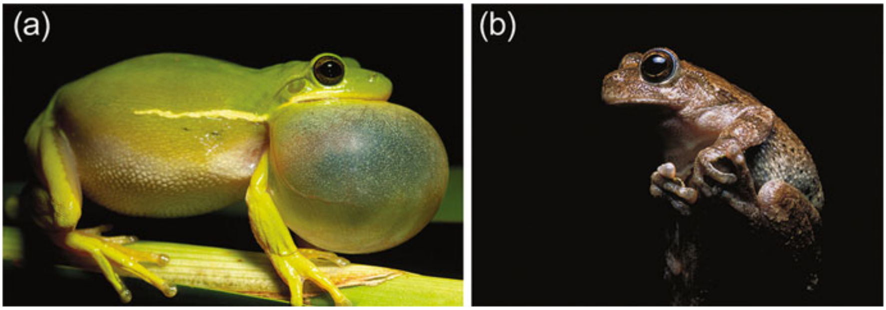


Fig. 4.4 Species differences in coloration and conspicuousness in hylid treefrogs. A calling male green treefrog, *Hyla cinerea* (*left*), and Cope's gray treefrog, *Hyla chrysoscelis* (*right*). Note that individuals of the gray treefrog are able to change their coloration from brown and marbled gray to a conspicuous solid green that resembles the green treefrog

related and exhibit similar ecologies, they express markedly different visual phenotypes (Fig. 4.4). Green treefrogs are vibrant green and possess a conspicuous lateral stripe (often much more pronounced than the example in Fig. 4.4). In contrast, gray treefrogs can vary in coloration from bright green to darker gray or brown, and they lack the prominent lateral stripe present in green treefrogs. The acoustic properties of male advertisement calls are also quite different; green treefrogs produce short (ca. 200 ms), rapidly repeated calls, whereas gray treefrogs produce a long, pulsed call (total call duration often exceeding 1 s) at much slower rates. In two-choice phonotaxis tests, Laird et al. (2016) demonstrated that female green treefrogs strongly prefer the visual stimulus of a robotic green treefrog (coupled to an advertisement call) over the same call absent the visual stimulus. Taylor et al. (2007) found that females of the closely related squirrel treefrog (*Hyla squirella*) preferred a static model frog with a large lateral stripe to one with a small stripe. Under typical nocturnal, low-illumination conditions, the stripe may simply increase visual contrast and provide a better target on which to group the acoustic stimulus. Although the role of the lateral stripe in the communication system of the green treefrog has not yet been examined, the stripe is even more pronounced than in squirrel treefrogs. This at least hints that a similar process may be occurring in the green treefrog.

In video playbacks, the eastern gray treefrog (*Hyla versicolor*) was shown to respond to visual stimuli of calling males (Reichert et al. 2014), but females were unselective in their responses to video manipulation of the vocal sac. In similar video playback experiments, Cope's gray treefrogs failed to discriminate between a static vocal sac and one that inflated with each presented call (K. M. Schrode and M. A. Bee unpublished data). What remains unclear, however, is whether the female gray treefrogs are truly unselective about characteristics of vocal sac movement (shown to be important in other species) or if the *phototactic* response of video screen lighting overwhelmed any important biological responses to visual stimuli. Although the picture is far from complete, the differences in acoustic mechanisms of auditory scene analysis between gray and green treefrogs, coupled with limited multisensory data, suggest that the two species may have evolved different sensory mechanisms to deal with the common problem of communicating in noisy environments.

4.3.2 Beyond Bimodal Grouping

A number of studies in humans have addressed the integration of senses other than audiovisual (Hotting and Roder 2004; Bresciani et al. 2005; Stevenson 2012). In nonhuman animals, several studies have also documented integration outside the audiovisual domain (Rundus et al. 2007; Leonard et al. 2011; Hebets and Rundus 2011; Clarke et al. 2013). To date however, the majority of studies have focused on the integration of acoustic and visual cues, especially in neurobiology (Stein 2012), and few studies have incorporated more than two modalities (Higham and Hebets 2013). This likely reflects human bias as well as the ease with which acoustic and visual signals can be experimentally manipulated. Recent studies have made it clear that animals also recruit information from additional modalities. Studies of multisensory integration outside the audiovisual domain are relatively limited, but represent an important next step in understanding perceptual grouping.

The results of two intriguing studies indicate that trimodal integration might be important in some frogs. The first is a study of African reed frogs (*Hyperoliidae*). As is typical of most frogs, male reed frogs inflate a conspicuous vocal sac while producing advertisement calls (Grafe 1996). Interestingly, these frogs possess a conspicuous patch on their vocal sac that, in many species, differs in color from the surrounding tissue (Fig. 4.5). Starnberger et al. (2013) conducted a detailed study of 11 species in 4 genera. They performed histological and chemical examinations of this gular patch tissue and found that the tissue formed secretory glands. They also identified 65 distinct compounds and found that the chemical profiles differed among species. Although Starnberger et al. (2013) did not conduct playback

Fig. 4.5 Trimodal signaling in frogs? Shown here is a calling male *Hyperolius cinnamomeoventris* (Hyperoliidae) with the yellow gular patch on the inflated vocal sac clearly visible. Photo courtesy I. Starnberger

studies, their data, coupled with limited information in other species (Waldman and Bishop 2004; Smith et al. 2004), strongly suggest that chemical cues may be involved with the communication systems of these frogs.

In túngara frogs, courting males establish calling sites at a pond and defend them against other males who approach too closely (typically to within 10 cm). This call site defense helps to space males and likely reduces acoustic competition with nearby neighbors. Males call while floating at the edge of the pond; the expansion of the vocal sac and subsequent abdominal movements necessarily generate ripples on the water surface (see Fig. 11.5 in Chap. 11). Ripples propagate at substantially slower speeds than airborne acoustic waves. Thus, for male receivers, arrival time differences accumulate with increasing distances between males. Halfwerk et al. (2014) showed that male túngara frogs conduct cross-modal comparisons by attending to these arrival time differences. In particular, they found that males increase their call rates (a response to competition from a nearby caller) more in response to ripples plus a call compared to a call alone. In addition, vocal response to simulated rivals decreased with increasing arrival time distances, demonstrating that the frogs are integrating acoustic and tactile cues. Some interesting questions remain. First, as distance between males increases, and hence arrival time differences accumulate, the arrival of the ripple from the first call may coincide with the arrival time from the second call. In this case, perceptual grouping may become increasingly difficult and males may have to rely on additional cues (Halfwerk et al. 2014). Also, interference between calls and ripples generated by multiple males likely generate cocktail-party-like problems in multiple modalities. In these situations, male receivers have a suite of cues they can use to develop object perceptions, including interactions among acoustic directionality, tactile directionality, arrival time differences, variation in acoustic amplitude, and visual cues. In túngara frogs, the data show that these animals recruit at least three modalities (acoustic, visual, and tactile) to assign signals to their source. How and when these components might be integrated as a trimodal signal remains to be investigated.

4.4 Summary and Future Directions

Like much of the literature on perceptual grouping in humans, this review took a stimulus-driven approach to explore how different stimulus parameters affect grouping. Drawing on examples from studies of anurans and orthopterans, we reviewed experiments showing that temporal, spectral, and spatial stimulus parameters are used to evaluate which stimuli are produced by the same source. Furthermore, with respect to auditory grouping, these parameters may be used in both primitive and schema-based perception of stimulus groups. Such perception is not limited to a single modality, however, with visual and auditory combination sensitivity clearly elucidated in these animals. Thus, the literature reviewed here supports the hypothesis that, across taxa, the importance of assigning stimuli to the correct source has resulted in similar perceptual solutions for stimulus grouping.

The extent to which there is also evolutionary homology in the underlying neural mechanisms of perceptual grouping across taxa remains an important and largely open question.

Although there has been progress, since Bregman's (1990) book, in developing animal models for experiments on perceptual grouping, much work remains. For example, we do not fully understand the neural code for scene analysis of real-world stimuli (Klump 2005; Chap. 3). Whereas extensive behavioral data and computational explanations can be found in the literature (Mellinger and Mont-Reynaud 1996; Wang and Brown 2006), data elucidating the cellular and systems based neural mechanisms underlying the grouping of real-world stimuli remain incomplete (Feng and Ratnam 2000; Feng and Schul 2007). This means that computational models explaining behavior, while based on many neural correlates, are partly phenomenological and, from a comparative point of view, do not offer insight into potential diversity of possible neural solutions that exist for this universal perceptual problem. More directly, without specifying neural codes for grouping, it is difficult to assess the extent of evolutionary homology and analogy in scene analysis. Of course, addressing such a question incorporates several others: what are the necessary and sufficient mechanisms for grouping; where in the circuitry do they exist; do specializations in some animals operate at any circuit level or are they limited to a particular node; does experimentally conditioned grouping behavior engage different mechanisms from those mediating functionally significant stimulus grouping; which mechanisms are innate and which are plastic; to what extent do descending mechanisms play a role in perceptual grouping; and where in the circuitry do they exercise their influence? This list of open questions is extensive and, intimidatingly, not exhaustive.

A fundamental question that must be operationally addressed in future work is, what are the neural responses that correspond to grouped versus ungrouped percepts? It is here that model organisms with well-defined grouping percepts, as revealed by the animal's behavioral decisions, may be most advantageous to building a framework for grouping mechanisms (Chap. 3). Whereas studies of communication in a diversity of animals have sought to elucidate how single components of signals are processed, animal models like the anurans reviewed here will be useful for discovering mechanisms for perceptually grouping multiple signal components. Several lines of research are ongoing. For example, neural adaptation is an important null hypothesis for the mechanisms of streaming (Naatanen and Alho 1995; Ulanovsky et al. 2004; Bee and Klump 2005; Schul and Sheridan 2006; Bee et al. 2010) and may be relevant to streaming the signals described above in frogs and crickets (Sobel and Tank 1994; Ponnath and Farris 2010; Ponnath et al. 2013). For neurons in the anuran midbrain, temporal characteristics of inhibitory and excitatory synaptic currents enable sequential counting of pulsed signals in frogs (Edwards et al. 2002; Rose et al. 2011). Furthermore, these midbrain and diencephalic neurons exhibit complex signal processing (Hoke et al. 2004, 2005, 2007), as well as AND-gate sensitivity (Mudry and Capranica 1987a, b). While such processing almost certainly reflects the properties of ascending circuitry, a complete understanding of grouping mechanisms must include descriptions and analyses of the contribution of descending circuitry. Our understanding of descending (attentional) mechanisms is largely incomplete in all taxa, especially with respect to the extent to which descending modulation is specific to stimulus parameters, an important factor in attention-driven perception of groups. So far, only a pair of studies has assessed the most basic effects of descending modulation on ascending auditory responses in frogs (Endepols and Walkowiak 2000; Ponnath and Farris 2014). Even from these basic studies, however, we know that midbrain auditory processing can be strongly modulated by thalamic and striatal input. Interestingly, this modulation can occur over short time scales (ms) and for sounds coming from different locations, potentially allowing for fast changes or biases in the processing of different sound sources. Such top-down modulation in frogs, when taken together with the data from ascending auditory units, suggest that neocortical involvement is not necessary for combination sensitivity and attentional control, a conclusion that was facilitated by a broad comparative approach. Nevertheless, it must be empirically determined to what extent the circuits of the mammalian neocortex and avian pallium add new sensory functions (including unimodal and cross-modal grouping) and build on those evolved and maintained in subcortical circuits (e.g., superior colliculus; Stein and Meredith 1993).

A full understanding of perceptual grouping, particularly including the neural responses of grouped versus ungrouped stimuli, requires a multistep approach. It requires first defining neural processing in a single modality within a model system. The next step is to compare this single sensory process to the bimodal and then possibly trimodal responses. Once these behavioral and neural responses are mapped in several different taxa, the homologous and analogous responses can then be examined in a comparative evolutionary framework. This is admittedly an ambitious undertaking, particularly since it involves addressing a variety of questions outlined earlier in this section, and because very little is known about multisensory integration beyond two modalities. Regardless, the past couple of decades have yielded substantial increases in our understanding of perceptual grouping both within and across modalities. The direction we reviewed here promises to provide a deep understanding of the universal phenomenon of perceptual grouping

Acknowledgments We thank Mark Bee and Kim Hunter for helpful discussions. We are grateful to Mike Ryan for his insight into mate choice behavior and for including us in “Team Túngara.” Iris Starnberger generously provided the photo of the calling African reed frog. Helpful resources were provided by N. Bazan, S. McClugage (LSUHSC), and K. Olmstead (SU), and funding was provided by the Smithsonian Tropical Research Institute and the National Science Foundation.

References

Aho AC, Donner K, Hyden C, Larsen LO, Reuter T (1988) Low retinal noise in animals with low body temperature allows high visual sensitivity. Nature 334(6180):348–350

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

Andersson M (1994) Sexual selection. Princeton University Press, Princeton

Bee MA (2010) Spectral preferences and the role of spatial coherence in simultaneous integration in gray treefrogs (*Hyla chrysoscelis*). J Comp Psychol 124:412–424

Bee MA (2015) Treefrogs as animal models for research on auditory scene analysis and the cocktail party problem. Int J Psychophysiol 95(2):216–237

Bee MA, Klump GM (2004) Primitive auditory stream segregation: a neurophysiological study in the songbird forebrain. J Neurophysiol 92(2):1088–1104

Bee MA, Klump GM (2005) Auditory stream segregation in the songbird forebrain: effects of time intervals on responses to interleaved tone sequences. Brain Behav Evol 66(3):197–214

Bee MA, Micheyl C (2008) The cocktail party problem: what is it? How can it be solved? And why should animal behaviorists study it? J Comp Psychol 122(3):235–251

Bee MA, Riemersma KK (2008) Does common spatial origin promote the auditory grouping of temporally separated signal elements in grey treefrogs? Anim Behav 76(3):831–843

Bee MA, Micheyl C, Oxenham AJ, Klump GM (2010) Neural adaptation to tone sequences in the songbird forebrain: patterns, determinants, and relation to the build-up of auditory streaming. J Comp Physiol A 196(8):543–557

Bennet-Clark HC (1989) Songs and the physics of sound production. In: Huber F, Moore TE, Loher W (eds) Cricket behavior and neurobiology. Cornell University Press, Ithaca, pp. 227–261

Bennet-Clark HC (1995) Insect sound production: transduction mechanisms and impedance matching. Symp Soc Exp Biol 49:199–218

Bennet-Clark HC (1999) Resonators in insect sound production: how insects produce loud pure-tone songs. J Exp Biol 202(Pt 23):3347–3357

Braaten RF, Hulse SH (1993) Perceptual organization of auditory temporal patterns in European starlings (*Sturnus vulgaris*). Percept Psychophys 54(5):567–578

Bradbury JW, Vehrencamp SL (1998) Principles of animal communication. Sinauer Association, Sunderland

Bregman AS (1990) Auditory scene analysis: the perceptual organization of sound. MIT Press, Cambridge

Bregman AS, Campbell J (1971) Primary auditory stream segregation and perception of order in rapid sequences of tones. J Exp Psychol 89(2):244–249

Bregman AS, Rudnicky AI (1975) Auditory segregation: stream or streams? J Exp Psychol Hum Percept Perform 1(3):263–267

Bresciani JP, Ernst MO, Drewing K, Bouyer G, Maury V, Kheddar A (2005) Feeling what you hear: auditory signals can modulate tactile tap perception. Exp Brain Res 162(2):172–180

Bronkhorst AW (2015) The cocktail-party problem revisited: early processing and selection of multi-talker speech. Atten Percept Psychophys 77(5):1465–1487

Buchanan BW (1993) Effects of enhanced lighting on the behaviour of nocturnal frogs. Anim Behav 45(5):893–899

Burke D, Everingham P, Rogers T, Hinton M, Hall-Aspland S (2001) Perceptual grouping in two visually reliant species: humans (*Homo sapiens*) and Australian sea lions (*Neophoca cinerea*). Perception 30(9):1093–1106

Bushara KO, Grafman J, Hallett M (2001) Neural correlates of auditory–visual stimulus onset asynchrony detection. J Neurosci 21(1):300–304

Capranica RR, Moffat AJM (1983) Neurobehavioral correlates of sound communication in anurans. In: Ewert JP, Capranica RR (eds) Andvances in vertebrate neuroethology. Plenum Press, New York, pp. 701–703

Cherry EC (1953) Some experiments on the recognition of speech, with one and two ears. J Acoust Soc Am 25:975–979

Cherry EC, Weary DM (1954) Some further experiments upon the recognition of speech, with one and with two ears. J Acoust Soc Am 26:554–559

Clarke D, Whitney H, Sutton G, Robert D (2013) Detection and learning of floral electric fields by bumblebees. Science 340(6128):66–69

Cummings ME, Bernal XE, Reynaga R, Rand AS, Ryan MJ (2008) Visual sensitivity to a conspicuous male cue varies by reproductive state in Physalaemus pustulosus females. J Exp Biol 211(8):1203–1210

Cusack R, Carlyon RP (2004) Auditory stream segregation inside and outside the laboratory. In: Neuhoff J (ed) Ecological psychoacoustics. Elsevier, San Diego, pp. 15–48

Dambach M, Gras A (1995) Bioacoustics of a miniature cricket, *Cycloptiloides canariensis* (Orthoptera: Gryllidae: Mogoplistinae). J Exp Biol 198(Pt 3):721–728

Darwin CJ (1997) Auditory grouping. Trends Cogn Sci 1(9):327–333

Darwin CJ, Carlyon RP (1995) Auditory grouping. In: Moore BC (ed) Hearing. Academic, San Diego, pp. 387–424

Darwin CJ, Brungart DS, Simpson BD (2003) Effects of fundamental frequency and vocal-tract length changes on attention to one of two simultaneous talkers. J Acoust Soc Am 114 (5):2913–2922

Dixon NF, Spitz L (1980) The detection of auditory visual desynchrony. Perception 9(6):719–721

Dooling RJ, Hulse SH (1989) The comparative psychology of audition: perceiving complex sounds. Lawrence Erlbaum Associates, Hillsdale

Edwards CJ, Alder TB, Rose GJ (2002) Auditory midbrain neurons that count. Nat Neurosci 5 (10):934–936

Elhilali M, Ma L, Micheyl C, Oxenham AJ, Shamma SA (2009) Temporal coherence in the perceptual organization and cortical representation of auditory scenes. Neuron 61(2):317–329

Elias DO, Hebets EA, Hoy RR (2006) Female preference for complex/novel signals in a spider. Behav Ecol 17(5):765–771

Endepols H, Walkowiak W (2000) Integration of ascending and descending inputs in the auditory midbrain of anurans. J Comp Physiol A 186(12):1119–1133

Farris HE and Ryan MJ (2011) Relative comparisons of call parameters enable auditory grouping in frogs. Nat Commun. doi:10.1038/NCOMMS1417

Farris HE, Rand AS, Ryan MJ (2002) The effects of spatially separated call components on phonotaxis in túngara frogs: evidence for auditory grouping. Brain Behav Evol 60(3):181–188

Farris HE, Rand AS, Ryan MJ (2005) The effects of time, space and spectrum on auditory grouping in túngara frogs. J Comp Physiol A 191(12):1173–1183

Fay RR (1988) Hearing in vertebrates: a psychophsyics databook. Hill-Fay Associates, Winnetka

Fay RR (2008) Sound source perception and stream segregation in nonhuman vertebrate animals. In: Yost WA, Popper AN, Fay RR (eds) Auditory perception of sound sources. Springer, New York, pp. 307–323

Fay RR, Popper AN (2000) Evolution of hearing in vertebrates: the inner ears and processing. Hear Res 149(1–2):1–10

Feng AS, Ratnam R (2000) Neural basis of hearing in real-world situations. Annu Rev Psychol 51:699–725

Feng AS, Schul J (2007) Sound processing in real-world environments. In: Narins PM, Feng AS, Fay RR, Popper AN (eds) Hearing and sound communication in amphibians. Springer, New York, pp. 323–350

Feng AS, Narins PM, Xu CH, Lin WY, Yu ZL, Qiu Q, Xu ZM, Shen JX (2006) Ultrasonic communication in frogs. Nature 440(7082):333–336

Fishman YI, Reser DH, Arezzo JC, Steinschneider M (2001) Neural correlates of auditory stream segregation in primary auditory cortex of the awake monkey. Hear Res 151(1–2):167–187

Fishman YI, Arezzo JC, Steinschneider M (2004) Auditory stream segregation in monkey auditory cortex: effects of frequency separation, presentation rate, and tone duration. J Acoust Soc Am 116(3):1656–1670

Forester DC, Czarnowsky R (1985) Sexual selection in the spring peeper, Hyla crucifer (Amphibia, Anura): role of the advertisement call. Behaviour 92(1):112–127

Freyman RL, Helfer KS, McCall DD, Clifton RK (1999) The role of perceived spatial separation in the unmasking of speech. J Acoust Soc Am 106(6):3578–3588

Freyman RL, Balakrishnan U, Helfer KS (2001) Spatial release from informational masking in speech recognition. J Acoust Soc Am 109(5 Pt 1):2112–2122

Gans C (1973) Sound production in the Salientia: mechanism and evolution of the emitter. Am Zool 13:1179–1194

Gaub S, Ehret G (2005) Grouping in auditory temporal perception and vocal production is mutually adapted: the case of wriggling calls of mice. J Comp Physiol A Neuroethol Sens Neural Behav Physiol 191(12):1131–1135

Gerhardt HC (1994) Reproductive character displacement of female mate choice in the grey treefrog. Anim Behav 47(4):959–969

Gerhardt HC (2001) Acoustic communication in two groups of closely related treefrogs. Adv Study Behav 30:99–167

Gerhardt HC, Bee MA (2007) Recognition and localization of acoustic signals. In: Narins PM, Feng AS, Fay RR, Popper AN (eds) Hearing and sound communication in amphibians, Springer handbook of auditory research, vol 28. Springer, New York, pp. 113–146

Gerhardt HC, Huber F (2002) Acoustic communication in insects and anurans. University of Chicago, Chicago

Grafe TU (1996) The function of call alternation in the African reed frog (*Hyperolius marmoratus*): precise call timing prevents auditory masking. Behav Ecol Sociobiol 38 (3):149–158

Gridi-Papp M, Rand AS, Ryan MJ (2006) Animal communication: complex call production in the túngara frog. Nature 441(7089):38

Gu JJ, Montealegre Z, Robert D, Engel MS, Qiao GX, Ren D (2012) Wing stridulation in a Jurassic katydid (Insecta, Orthoptera) produced low-pitched musical calls to attract females. Proc Natl Acad Sci USA 109(10):3868–3873

Halfwerk W, Page RA, Taylor RC, Wilson PS, Ryan MJ (2014) Crossmodal comparisons of signal components allow for relative-distance assessment. Curr Biol 24(15):1751–1755

Hartmann WM, Johnson D (1991) Stream segregation and peripheral channeling. Music Percept 9:155–184

Hawley ML, Litovsky RY, Culling JF (2004) The benefit of binaural hearing in a cocktail party: effect of location and type of interferer. J Acoust Soc Am 115(2):833–843

Hebets EA (2008) Seismic signal dominance in the multimodal courtship display of the wolf spider Schizocosa stridulans Stratton 1991. Behav Ecol 19(6):1250–1257

Hebets EA and Rundus A (2011) Chemical communication in a multimodal context. Chemical Communication in Crustaceans. Springer, pp 335–354

Henson RA (1977) Neurological aspects of musical experience. In: Henson RA (ed) Music and the brain: studies in the neurology of music. Charles C Thomas, Springfield

Herberstein ME, Wignall AE, Hebets EA, Schneider JM (2014) Dangerous mating systems: signal complexity, signal content and neural capacity in spiders. Neurosci Biobehav Rev 46 (Pt 4):509–518

Higham JP, Hebets EA (2013) An introduction to multimodal communication. Behav Ecol Sociobiol 67(9):1381–1388

Hillery CM (1984) Seasonality of two midbrain auditory responses in the treefrog, *Hyla chrysoscelis*. Copeia 1984(4):844–852

Hirschmann W, Hödl W (2006) Visual signaling in Phrynobatrachus krefftii boulenger, 1909 (Anura: Ranidae). Herpetologica 62(1):18–27

Hoke KL, Burmeister SS, Fernald RD, Rand AS, Ryan MJ, Wilczynski W (2004) Functional mapping of the auditory midbrain during mate call reception. J Neurosci 24(50):11264–11272

Hoke KL, Ryan MJ, Wilczynski W (2005) Social cues shift functional connectivity in the hypothalamus. Proc Natl Acad Sci USA 102(30):10712–10717

Hoke KL, Ryan MJ, Wilczynski W (2007) Integration of sensory and motor processing underlying social behaviour in túngara frogs. Proc Biol Sci 274(1610):641–649

Hotting K, Roder B (2004) Hearing cheats touch, but less in congenitally blind than in sighted individuals. Psychol Sci 15(1):60–64

Howard IP, Templeton WB (1966) Human spatial orientation. Wiley, London

Hulse SH (2002) Auditory scene analysis in animal communication. In: Slater P, Rosenblatt J, Snowdon C, Roper T (eds) Advances in the study of behavior, vol 31. Academic, New York, pp. 163–200

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

Klump G (2005) How does the hearing system perform auditory scene analysis? In: van Hemmen JL, Sejnowski TJ (eds) Problems in systems neuroscience. Oxford University Press, New York, pp. 303–321

Knudsen EI (1982) Auditory and visual maps of space in the optic tectum of the owl. J Neurosci 2 (9):1177–1194

Knudsen EI, Knudsen PF (1989) Vision calibrates sound localization in developing barn owls. J Neurosci 9(9):3306–3313

Kozak EC, Uetz GW (2016) Cross-modal integration of multimodal courtship signals in a wolf spider. Anim Cogn. doi:10.1007/s10071-016-1025-y

Kurylo DD, Van NJ, Knepper B (1997) Characteristics of perceptual grouping in rats. J Comp Psychol 111(2):126–134

Laird KL, Clements P, Hunter KL, Taylor RC (2016) Multimodal signaling improves mating success in the green tree frog (*Hyla cinerea*), but may not help small males. Behav Ecol Sociobiol 70:1517–1525

Larson KA (2004) Advertisement call complexity in Northern Leopard Frogs, *Rana pipiens*. Copeia 3:676–682

Leonard AS, Dornhaus A, Papaj DR (2011) Flowers help bees cope with uncertainty: signal detection and the function of floral complexity. J Exp Biol 214(1):113–121

Lindquist ED, Hetherington TE (1998) Semaphoring in an earless frog: the origin of a novel visual signal. Anim Cogn 1(2):83–87

Lombardo SR, Mackey E, Tang L, Blumstein DT (2008) Multimodal communciation and spatial binding in pied currawongs (*Strepera graculina*). Anim Cogn 11(4):675–682

MacDougall-Shackleton SA, Hulse SH, Gentner TQ, White W (1998) Auditory scene analysis by European starlings (Sturnus vulgaris): perceptual segregation of tone sequences. J Acoust Soc Am 103(6):3581–3587

Martin WF (1971) Mecanisms of sound production in toads of the genus *Bufo*: passive elements. J Exp Zool 176:273–294

Maruska KP, Ung US, Fernald RD (2012) The African cichlid fish Astatotilapia burtoni uses acoustic communication for reproduction: sound production, hearing, and behavioral significance. PLoS One 7(5):e37612

Maynard Smith J, Harper DGC (2003) Animal signals. Oxford University Press, Oxford

McGurk H, Macdonald J (1976) Hearing lips and seeing voices. Nature 264(5588):746–748

Mellinger DK, Mont-Reynaud BM (1996) Scene analysis. In: Popper AN, Fay RR (eds) Auditory computation. Springer, Berlin, pp. 271–331

Micheyl C, Shamma S, Elhilali M, Oxenham AJ (2010). Sequential and simultaneous auditory grouping measured with synchrony detection. In: EA Lopez-Poveda, AR Palmer, R Meddis, eds The neurophysiological bases of auditory perception, Springer, New York, pp. 489-496

Miller GA, Heise GA (1950) The trill threshold. J Acoust Soc Am 22(5):637–638

Moore BCJ (2012) An introduction to the psychology of hearing, 6th edn. Emeral Group, Bingley

Moore BCJ, Gockel H (2002) Factors influencing sequential stream segregation. Acta Acust United Ac 88(3):320–333

Moore BC, Gockel HE (2012) Properties of auditory stream formation. Philos Trans R Soc Lond Ser B Biol Sci 367(1591):919–931

Mudry KM, Capranica RR (1987a) Correlation between auditory evoked responses in the thalamus and species-specific call characteristics. I. *Rana catesbeiana* (Anura: Ranidae). J Comp Physiol A 160(4):477–489

Mudry KM, Capranica RR (1987b) Correlation between auditory thalamic area evoked responses and species-specific call characteristics. II. H. *Hyla cinerea* (Anura: Hylidae). J Comp Physiol A 161(3):407–416

Munoz NE and Blumstein DT (2012) Multisensory perception in uncertain environments. Behav Ecol:arr220

Naatanen R, Alho K (1995) Mismatch negativity – a unique measure of sensory processing in audition. Int J Neurosci 80(1–4):317–337

Narins PM, Hödl W, Grabul DS (2003) Bimodal signal requisite for agonistic behavior in a dart-poison frog, *Epipedobates femoralis*. Proc Natl Acad Sci 100(2):577–580

Narins PM, Grabul DS, Soma KK, Gaucher P, Hödl W (2005) Cross-modal integration in a dart-poison frog. Proc Natl Acad Sci USA 102(7):2425–2429

Neiworth JJ, Whillock KM, Kim SH, Greenberg JR, Jones KB, Patel AR, Steefel-Moore DL, Shaw AJ, Rupert DD, Gauer JL, Kudura AG (2014) Gestalt principle use in college students, children with autism, toddlers (*Homo sapiens*), and cotton top tamarins (*Saguinus oedipus*). J Comp Psychol 128(2):188–198

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(6):e21191

Nityananda V, Bee MA (2012) Spatial release from masking in a free-field source identification task by gray treefrogs. Hear Res 285(1):86–97

Partan SR, Larco CP, Owens MJ (2009) Wild tree squirrels respond with multisensory enhancement to conspecific robot alarm behaviour. Anim Behav 77(5):1127–1135

Pollack GS, Huber F, Weber T (1984) Frequency and temporal pattern-dependent phonotaxis of crickets (*Teleogryllus oceanicus*) during tethered flight and compensated walking. J Comp Physiol 154(1):13–26

Ponnath A, Farris HE (2010) Calcium-dependent control of temporal processing in an auditory interneuron: a computational analysis. J Comp Physiol A Neuroethol Sens Neural Behav Physiol 196(9):613–628

Ponnath A, Farris HE (2014) Sound-by-sound thalamic stimulation modulates midbrain auditory excitability and relative binaural sensitivity in frogs. Front Neural Circuits 8:85

Ponnath A, Hoke KL, Farris HE (2013) Stimulus change detection in phasic auditory units in the frog midbrain: frequency and ear specific adaptation. J Comp Physiol A 199(4):295–313

Popper AN, Fay RR (1997) Evolution of the ear and hearing: issues and questions. Brain Behav Evol 50(4):213–221

Preininger D, Boeckle M, Freudmann A, Starnberger I, Sztatecsny M, Hödl W (2013a) Multimodal signaling in the Small Torrent Frog (Micrixalus saxicola) in a complex acoustic environment. Behav Ecol Sociobiol 67(9):1449–1456

Preininger D, Boeckle M, Sztatecsny M, Hödl W (2013b) Divergent receiver responses to components of multimodal signals in two foot-flagging frog species. PLoS One 8(1):e55367

Pressnitzer D, Sayles M, Micheyl C, Winter IM (2008) Perceptual organization of sound begins in the auditory periphery. Curr Biol 18(15):1124–1128

Reichert MS, Galante H, Höbel G (2014) Female gray treefrogs, Hyla versicolor, are responsive to visual stimuli but unselective of stimulus characteristics. J Exp Biol 217(18):3254–3262

Reichert MS, Symes LB, Höbel G (2016) Lighting up sound preferences: cross-modal influences on the precedence effect in treefrogs. Anim Behav 119:151–159

Rock I, Brosgole L (1964) Grouping based on phenomenal proximity. J Exp Psychol 67 (6):531–538

Rock I, Palmer S (1990) The legacy of Gestalt psychology. Sci Am 263(6):84–90

Rose MM, Moore BC (2000) Effects of frequency and level on auditory stream segregation. J Acoust Soc Am 108(3 Pt 1):1209–1214

Rose GJ, Leary CJ, Edwards CJ (2011) Interval-counting neurons in the anuran auditory midbrain: factors underlying diversity of interval tuning. J Comp Physiol A 197(1):97–108

Rosenthal GG, Rand AS, Ryan MJ (2004) The vocal sac as a visual cue in anuran communication: an experimental analysis using video playback. Anim Behav 68(1):55–58

Rundus AS, Owings DH, Joshi SS, Chinn E, Giannini N (2007) Ground squirrels use an infrared signal to deter rattlesnake predation. Proc Natl Acad Sci 104(36):14372–14376

Ryan MJ (1985) The Túngara frog, a study in sexual selection and communication. University of Chicago Press, Chicago

Ryan MJ, Drewes RC (1990) Vocal morphology of the *Physalaemus-pustulosus* species group (Leptodactylidae) – morphological response to sexual selection for complex calls. Biol J Linn Soc Lond 40(1):37–52

Ryan MJ, Rand AS (1990) The sensory basis of sexual selection for complex calls in the Túngara frog, *Physalaemus pustulosus* (Sexual selection for sensory exploitation). Evolution 44 (2):305–314

Ryan MJ, Fox JH, Wilczynski W, Rand AS (1990) Sexual selection for sensory exploitation in the frog *Physalaemus pustulosus*. Nature 343(6253):66–67

Schildberger K (1984) Temporal selectivity of identified auditory neurons in the cricket brain. J Comp Physiol 155(2):171–185

Schrode KM, Buerkle NP, Brittan-Powell EF, Bee MA (2014) Auditory brainstem responses in Cope’s gray treefrog (*Hyla chrysoscelis*): effects of frequency, level, sex and size. J Comp Physiol A 200(3):221–238

Schul J, Sheridan RA (2006) Auditory stream segregation in an insect. Neuroscience 138(1):1–4

Schwartz JJ, Gerhardt HC (1995) Directionality of the auditory system and call pattern recognition during acoustic interference in the gray tree frog, *Hyla versicolor*. Audit Neurosci 1:195–206

Shaw KL, Herlihy DP (2000) Acoustic preference functions and song variability in the Hawaiian cricket Laupala cerasina. Proc Biol Sci 267(1443):577–584

Simmons DD, Meenderink SWF, Vassilakis PN (2007) Anatomy, physiology, and function of auditory end-organs in the frog inner ear. In: Narins PM, Feng AS, Fay RR, Popper AN (eds) Hearing and sound communication in amphibians. Springer, New York, pp. 184–220

Smith BP, Williams CR, Tyler MJ, Williams BD (2004) A survey of frog odorous secretions, their possible functions and phylogenetic significance. Appl Herpetol 2(1):47–82

Sobel EC, Tank DW (1994) In-vivo Ca2+ dynamics in a cricket auditory neuron – an example of chemical computation. Science 263(5148):823–826

Spinozzi G, De LC, Castelli S (2004) Detection of grouped and ungrouped parts in visual patterns by tufted capuchin monkeys (*Cebus apella*) and humans (*Homo sapiens*). J Comp Psychol 118 (3):297–308

Spinozzi G, De LC, Truppa V, Castorina G (2009) The relative use of proximity, shape similarity, and orientation as visual perceptual grouping cues in tufted capuchin monkeys (*Cebus apella*) and humans (*Homo sapiens*). J Comp Psychol 123(1):56–68

Starnberger I, Poth D, Peram PS, Schulz S, Vences M, Knudsen J, Barej MF, Rodel M-O, Lzl M, Odl W (2013) Take time to smell the frogs: vocal sac glands of reed frogs (Anura: Hyperoliidae) contain species-specific chemical cocktails. Biol J Linn Soc Lond 110 (4):828–838

Starnberger I, Preininger D, Hödl W (2014a) From uni-to multimodality: towards an integrative view on anuran communication. J Comp Physiol A 200(9):777–787

Starnberger I, Preininger D, Hödl W (2014b) The anuran vocal sac: a tool for multimodal signalling. Anim Behav 97:281–288

Stein BE (2012) The new handbook of multisensory processing. MIT Press, Cambridge

Stein BE, Meredith A (1993) The merging of the senses. MIT Press, Cambridge

Stevenson R (2012) Multisensory interactions in flavor perception. In: Stein BE (ed) The new handbook of multisensory processing. MIT Press, Cambridge

Sztatecsny M, Preininger D, Freudmann A, Loretto MC, Maier F, Hödl W (2012) Don’t get the blues: conspicuous nuptial colouration of male moor frogs (*Rana arvalis*) supports visual mate recognition during scramble competition in large breeding aggregations. Behav Ecol Sociobiol 66(12):1587–1593

Taylor RC, Ryan MJ (2013) Interactions of multisensory components perceptually rescue túngara frog mating signals. Science 341(6143):273–274

Taylor RC, Buchanan BW, Doherty JL (2007) Sexual selection in the squirrel treefrog *Hyla squirella*: the role of multimodal cue assessment in female choice. Anim Behav 74 (6):1753–1763

Taylor RC, Klein BA, Stein J, Ryan MJ (2008) Faux frogs: multimodal signalling and the value of robotics in animal behaviour. Anim Behav 76(3):1089–1097

Taylor RC, Klein BA, Stein J, Ryan MJ (2011) Multimodal signal variation in space and time: how important is matching a signal with its signaler? J Exp Biol 214(5):815–820

Ulanovsky N, Las L, Farkas D, Nelken I (2004) Multiple time scales of adaptation in auditory cortex neurons. J Neurosci 24(46):10440–10453

Uy JA, Safran RJ (2013) Variation in the temporal and spatial use of signals and its implications for multimodal communication. Behav Ecol Sociobiol 67(9):1499–1511

van Noorden LP (1977) Minimun differences of level and frequency for perceptual fission of tone sequences ABAB. J Acoust Soc Am 61(4):1041–1045

Vélez A, Bee MA (2011) Dip listening and the cocktail party problem in grey treefrogs: signal recognition in temporally fluctuating noise. Anim Behav 82(6):1319–1327

Vélez A, Hobel G, Gordon NM, Bee MA (2012) Dip listening or modulation masking? Call recognition by green treefrogs (Hyla cinerea) in temporally fluctuating noise. J Comp Physiol A Neuroethol Sens Neural Behav Physiol 198(12):891–904

Waldman B, Bishop PJ (2004) Chemical communication in an archaic anuran amphibian. Behav Ecol 15(1):88–93

Walker TJ (1957) Specificity in the response of female tree crickets (Orthoptera: Gryllidae: Oecanthinae) to calling songs of the males. Ann Entomol Soc Am 50:626–636

Walker TJ (1962) Factors responsible for intraspecific variation in calling songs of crickets. Evolution 16(4):407–428

Walker TJ (1974) Character displacement and acoustic insects. Am Zool 14(4):1137–1150

Wang D, Brown GJ (2006) Comutational auditory scene analysis: principles, algorithms, and applications. Wiley, Hoboken

Weber T, Thorson J (1988) Auditory-behavior of the cricket. 4. Interaction of direction of tracking with perceived temporal pattern in split-song paradigms. J Comp Physiol A 163(1):13–22

Wells KD (1980) Social behavior and communication of a dendrobatid frog (Colostethus trinitatis). Herpetologica 36:189–199

Whitchurch EA, Takahashi TT (2006) Combined auditory and visual stimuli facilitate head saccades in the barn owl (*Tyto alba*). J Neurophysiol 96(2):730–745

Yost WA, Sheft S (1993) Auditory perception. In: Yost WA, Popper AN, Fay RR (eds) Human psychophysics. Springer, New York, pp. 193–236

Yost WA, Popper AN, Fay RR (2008) Auditory perception of sound sources. Springer, New York

Chapter 5
Why Complex Signals Matter, Sometimes

Tricia L. Rubi and David W. Stephens

Abstract Animal signals commonly consist of multiple components—say a sound and a display—and students of signaling have offered many perceptual and cognitive explanations for why compound signals should be more effective. Yet, the economic benefits that receivers obtain by following multiple signal components remain unclear. Superficially, it would seem that a single discriminable difference should be sufficient to discriminate between underlying states, such as high-quality versus low-quality mates. This chapter asks when receivers can benefit by responding to combinations of signals. While there are many situations in which it is best to follow the single most reliable signal and ignore others, our model suggests that it can pay to follow signal combinations when these combinations indicate the occurrence of a rare event. This chapter develops the logic of this *confirmation of rare events* hypothesis of multiple signal use and discusses the implications of this idea for future studies of signaling.

5.1 Introduction

Drosophila males woo prospective mates with an elaborate mating display that includes singing, dancing, and tapping and licking the female. Eland bulls broadcast their fighting ability via a prominent dewlap, a tuft of dark facial hair, and a loud clicking produced by a tendon in the knee joint. Barn swallow nestlings beg for food by thrusting their bodies upward, vocalizing, and displaying a colorful gape that reflects both visible and ultraviolet wavelengths. Investigators have observed

---

T.L. Rubi (✉)
Department of Psychology, University of Michigan, Ann Arbor, MI, USA
e-mail: tricia.rubi@gmail.com

D.W. Stephens
Department of Ecology, Evolution, and Behavior, University of Minnesota, Twin Cities,
St. Paul, MN, USA
e-mail: dws@umn.edu

M.A. Bee, C.T. Miller (eds.), *Psychological Mechanisms in Animal Communication*, Animal Signals and Communication 5,
DOI 10.1007/978-3-319-48690-1_5

multicomponent signals, also known as complex signals, across a wide range of taxa and in every major type of signaling interaction, including courtship (reviewed in Candolin 2003), aposematism (reviewed in Pearson 1989; Rowe and Halpin 2013), agonism (e.g., Deag and Scott 1999; Chap. 4), begging displays (e.g., Leonard et al. 2003; Kim et al. 2011), sex identification (Page and Jaeger 2004), conspecific identification (de Caprona and Ryan 1990), and eavesdropping (Chap. 11). Signal components may be discrete qualities, such as the song and plumage of a courting bird or the color and odor of a baboon sexual swelling. Alternatively, components may be more tightly integrated, such as the structural and pigmentary properties of a single color patch (Grether et al. 2004) or lip movements and speech sounds in human verbal communication (Massaro and Cohen 1990). As our ability to observe and quantify traits improves, an increasing number of communication researchers argue that most signals consist of multiple components (e.g., Partan and Marler 1999; Rowe and Skelhorn 2004; Hebets and Papaj 2005; Wilson et al. 2013).

A large body of work indicates that additional signal components can increase the effectiveness of signals via a range of perceptual and cognitive mechanisms. Evidence shows, for example, that additional components can improve both the probability and speed of detection (reviewed in Rowe 1999; Hebets and Papaj 2005). Moreover, noisy environments can amplify these effects, especially if different sources of environmental noise have different effects on signal components (Candolin 2003; Hebets and Papaj 2005; Wilson et al. 2013). Some investigators have suggested that additional signal components may improve receiver performance through attentional effects. In the phenomenon of alerting, for example, a conspicuous component draws a receiver's attention to an informative component (Hebets and Papaj 2005). Multicomponent signals may also be learned faster and remembered longer (reviewed in Rowe 1999; Hebets and Papaj 2005).

The focus on perceptual and cognitive effects in the complex signaling literature is due, in part, to the fact that economic models of communication rarely find that following multiple components can be a viable receiver strategy. From a purely informational perspective, a single discriminable difference is sufficient to distinguish between two underlying states or conditions. While some authors have outlined specialized conditions that can favor complex signaling (e.g., specific cost structures or multiple receiver types), economic models have repeatedly found that for a basic signal (i.e., a redundant, reliable, cost-free signal), multiple components are no better than one component (Schluter and Price 1993; Johnstone 1995; Bro-Jørgensen 2010; Wilson et al. 2013; Dunlap and Stephens 2014; Rubi and Stephens 2016). Complex signals are prevalent in nature, suggesting that they are broadly beneficial across diverse signaling systems. Therefore, this is a surprising result even given the perceptual and cognitive benefits described above. Economic benefits and perceptual and cognitive benefits are often presented as alternatives; however, they are not mutually exclusive. Indeed, one might expect these effects to be intimately related; natural and sexual selection arising from communication can shape perceptual and cognitive mechanisms, and perceptual and cognitive mechanisms can impose selection on signals. Identifying the economic benefits of complex signaling may help us better understand the perceptual and cognitive mechanisms used by receivers to process multiple components.

This chapter presents a model that identifies economic situations in which receivers benefit from attending to combinations of signal components. We first review a "single-signal component" model (i.e., the flag model, McLinn and Stephens 2006; Dunlap and Stephens 2009; McLinn and Stephens 2010). We then build on this model by adding a second independent signal component. A key step in this model is to rigorously specify what it means to follow a combination of signal components. In agreement with previous economic models, our results show that a single signal component is often sufficient. However, we also identify a specific situation in which a receiver can economically benefit from attending to signal combinations. We argue that situations of this type may be reasonably common in nature.

5.2 The Model

5.2.1 *Alternative Actions, Uncertainty About Actions, Payoffs, and Signals*

Consider an animal that must choose between two actions. We call these options accept and reject, although we could equally well call them actions a and b. We suppose that uncertainty exists about which action is best, and we represent this uncertainty via a simple Bernoulli probability. The accept action is best with probability p, and we say that the "good condition" holds; it follows that the reject action is best with probability $(1-p)$, and we say that the "bad condition" holds. (Both of these probabilities provide a measure of the animal's uncertainty about the best action—hereafter, we use the terms p and "environmental uncertainty" interchangeably.) We consider a simple payoff structure in which the animal obtains one unit of benefit if it correctly matches its action to the environmental condition (i.e., if it accepts in the good condition or it rejects in the bad condition), and similarly it obtains zero units of benefit if it chooses the incorrect action (i.e., if it accepts in the bad condition or it rejects in the good condition). The animal can observe two forms of experience (or signals) that we call S and T. Each of the two signals can take one of two forms: S can take forms S+ or S−, and T can take forms T+ or T−. These are signals in the sense that the states of S and T are statistically related to the underlying condition (i.e., the good or accept-is-best condition and the bad or reject-is-best condition).

5.2.2 *Reliabilities*

We define the variable q such that $\mathrm{P(S+|\,Good)} = \mathrm{P(S-|\,Bad)} = q$. The variable q represents the reliability of S because if $q = 1.0$, then S is a perfect indicator of the underlying condition, but if $q = 0.5$ then S carries no information about the

Table 5.1 Variables and their definitions

Variable	Definition
p	Probability that the true condition is “Good”/ Environmental uncertainty
$1 - p$	Probability that the true condition is “Bad”
q	Reliability of signal S (more reliable signal)
r	Reliability of signal T (less reliable signal)
S+	Form of signal S indicating that the true condition is “Good”
S−	Form of signal S indicating that the true condition is “Bad”
T+	Form of signal T indicating that the true condition is “Good”
T−	Form of signal T indicating that the true condition is “Bad”

underlying condition. The model does not allow q values lower than 0.5, although they are perfectly reasonable biologically. We exclude them as a matter of definition because if q is less that 0.5 we simply can relabel the states of the signal so that S− becomes S+ and S+ becomes S−, which has the effect of keeping q in the range of 0.5–1.0. Now we define a similar variable r to represent the reliability of the second signal T so that $P(T+|\text{Good}) = P(T-|\text{Bad}) = r$. Again, we restrict r to the range 0.5–1.0. In addition, we assume that $q \geq r$, so that we can be sure that S is the more reliable of the two signals. All variables and their definitions are listed in Table 5.1.

5.2.3 Preliminaries: When Should a Single Signal Be Followed?

Under what conditions does it pay for a receiver to attend to both S and T? This is actually a somewhat more difficult and subtle question than it appears. To begin, it is important to understand the predicted use for a single signal, S. By signal use, we mean following a rule of the form accept when S+ is observed and reject when S− is observed. The alternative to signal use is, obviously enough, to choose the same action regardless of the observed state of the signal. So we must compare the value of signal use to the value of two alternative strategies: namely, always accept and always reject. Since the variable p gives us the relative likelihood of the good and bad conditions, we can easily calculate the payoffs associated with the three strategies of interest.

We can readily calculate the expected value of “always accept” using the following logic. The good condition occurs with probability p, and in the good condition accepting yields one unit; the bad condition occurs with probability $(1 - p)$, and in the bad condition accepting yields zero. So we have:

$$p \times 1 + (1 - p) \times 0 = p$$

To calculate the expected value of “always reject,” we follow the same type of reasoning, and we find:

$$p \times 0 + (1 - p) \times 1 = 1 - p$$

Next, we want to calculate the expected value of following the signal. Recall that the signal S matches the environmental condition with reliability q, and the "follow signal" strategy is defined to be "accept if S+ and reject if S−." When the condition is good, the signal shows S+ with probability q and S− with probability $(1 - q)$. So the signal follower will accept with probability q, gaining one unit, and reject with probability $(1 - q)$, gaining zero. Hence, a signal follower expects to gain $q \times 1 + (1 - q) \times 0$ in the good condition. Similarly, in the bad condition the signal shows state S+ with probability $(1 - q)$ and accepting yields zero; the signal shows state S− with probability q and rejecting yields 1 unit of benefit. Therefore, our signal follower expects to gain $(1 - q) \times 0 + q \times 1$ in the bad condition. Since the good condition occurs with probability p and the bad condition occurs with probability $(1 - p)$, the signal follower's overall expected gains are:

$$p \times [q \times 1 + (1 - q) \times 0] + (1 - p) \times [(1 - q) \times 0 + q \times 1] = q$$

Now we want to know which of these three strategies yields the highest expected benefit, and perhaps the easiest way to see this is to plot the three expected payoffs as functions of p (Fig. 5.1).

One can see that at the extreme values of p, it is better to ignore the signal. Specifically, if p is small it is better to ignore the signal and always reject because the fact that p is small means the bad condition is quite common and rejecting is the best single action. On the other hand, if p is large it is better to ignore the signal and always accept because the fact that p is large means that the good condition is quite common and accepting is the best single action. However, at intermediate values of p, when conditions are neither predominantly good nor bad, it is better to use the signal.

To summarize, for the single source of information case, we divide the p-axis into three regions: (1) for small p ($p < 1 - q$), it is best to ignore the signal and always reject; (2) for intermediate p ($1 - q < p < q$), it is best to follow the signals; and (3) for large p ($p > q$), it is best to ignore the signal and always accept. Some readers may recognize this as the so-called flag model (Dunlap and Stephens 2009; McLinn and Stephens 2010). This simple model is actually very rich and useful in its own right, but the goal of this paper is to consider economic benefits of using two signals in combination.

5.3 Following Two Signals

It may seem that the next step should be obvious. We already know the expected payoffs associated with using a single signal, so we simply want to compare this to the expected payoff associated with using two. However, it is not immediately obvious what it means to use two signals. Does it mean to accept if and only if S+

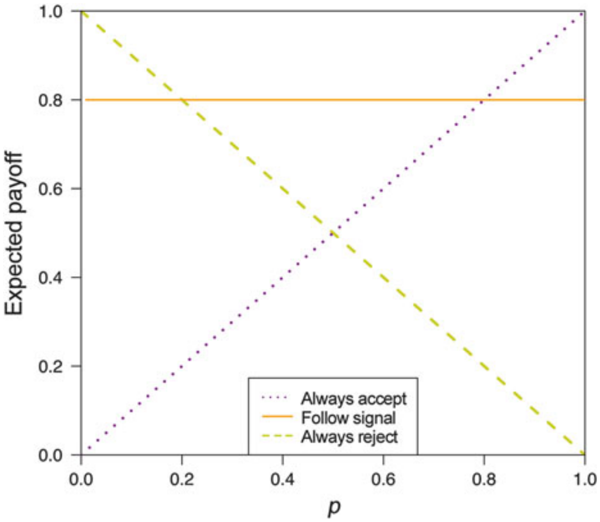


Fig. 5.1 When should a receiver follow a single signal? This plot shows expected values for three simple strategies as a function of p (on the x-axis): the increasing line shows the value of always accepting (p); the decreasing line shows the value of always rejecting ($1 - p$); and the horizontal line of height q gives the value of changing behavior in response to the signal S

and T+ are observed, or does it mean to always accept when S+ is observed and accept at some reduced probability when T+ is observed? In fact, a two-signal rule could, in principle, take many possible forms. Consider Table 5.2, where a, b, c, and d represent different probabilities of "accepting" for the four possible combinations of signal states. Using this table, it is very clear what it means to follow only one signal. For a decision maker who follows only S (the most reliable of the two signals by assumption), we would set $a = c = 1$ and $b = d = 0$. However, it is much less clear what it means to follow two signals. We could infer that following two simply means not following one, so any values of a, b, c, and d other than $a = c = 1$, $b = d = 0$, could reasonably be called following two signals.

It is obvious, we think, that to make progress, we must focus our attention on a reasonable subset of possible "two-signal" rules. To achieve this, we present the following argument. First, we consider only rules in which a, b, c, and d are set to either zero or one[1]. Second, we restrict our attention to rules in which $a = 1$ and $d = 0$. The argument here is that any rule that can reasonably be called a signal-following rule should accept when "plus–plus" combinations occur and reject when the "minus–minus" combination occurs. Recall that the "plus" states of the two signals are, by definition, statistically associated with the good condition; so, for example, a rule that set $a = 0$ and $d = 1$ would be a sort of anti-following rule, and would not make sense. Third, we can now restrict our attention to four possible "two-signal" rules that depend on the values of the anti-diagonal elements b and c. We consider them in turn: $b = 1$, $c = 0$, corresponds to follow only S rule discussed above, so we do not need to consider it as a "two-signal rule"; similarly, if $c = 1$ and

---

[1] There is no need to consider other values because the expected benefit of any rule specified by the acceptance probabilities a, b, c, and d will be a linear function of a, b, c, and d, and the maximal benefit will, therefore, occur at the extreme values of one or zero.

Table 5.2 Probabilities of accepting each of the four possible combinations of signal states for signal S and signal T. These probabilities (*a, b, c,* and *d*) are used to outline the signal-following rules

Signal T	Signal S	
	S+	S−
T+	*a*	*b*
T−	*c*	*d*

$b = 0$ then we have to follow only T rule which again is not a "two-signal rule." This leaves us with two plausible two-signal rules.

Two-signal rule #1. The *preponderance of positive evidence rule*: accept if and only if S+ and T+ are observed ($a = 1, \mathrm{b} = c = d = 0$)

Two-signal rule #2. The *preponderance of negative evidence rule*: reject if and only if S− and T− are observed ($a = b = c, d = 0$).

Now we can make progress. We can calculate the expected payoff derived from following each of two "two-signal rules" and superimpose the results on our "single-signal" analysis. So, in this new, larger analysis, we will compare the expected payoffs of five strategies. These five strategies are: (1) ignore all signals and always reject; (2) the preponderance of positive evidence rule (accept only when both S+ and T+ are observed); (3) follow the single most reliable signal (accept if S+, reject if S−); (4) the preponderance of negative evidence rule (reject only when both S− and T− are observed); and (5) ignore all signals and always accept.

To make specific algebraic predictions about these five possibilities, we must first specify the expected benefits associated with the two "new" two-signal rules. To calculate these expected payoffs, we follow the same basic logic as in our calculation of the expected benefits from following a single signal. Specifically, we first calculate the expected payoffs to a "two-signal user" given that the condition is good, then calculate the expected payoffs given that the condition is bad, and finally we combine these using the fact that the parameter p specifies the relative frequency of the good and bad conditions. The expected payoff from the preponderance of positive evidence rule is:

$$pqr + (1 - p)[1 - (1 - q)(1 - r)]$$

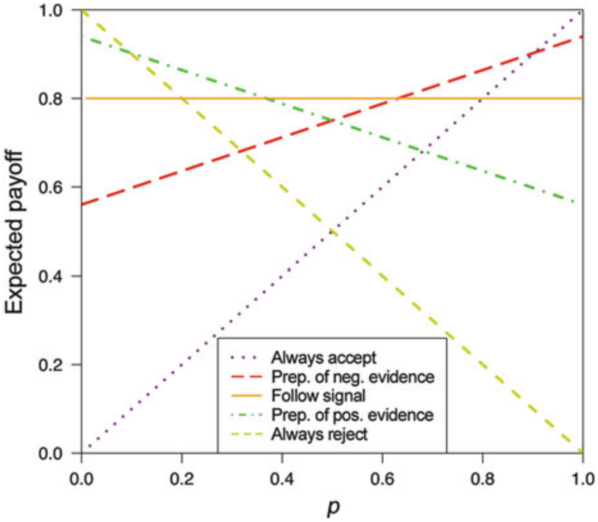
and the expected payoff from the preponderance of negative evidence rules is:

$$p[1 - (1 - q)(1 - r)] + (1 - p)qr$$

Figure 5.2 plots these two terms as functions of p and superimposes the results on our single-signal plot (shown in Fig. 5.1) to find a plot that includes the two two-signal rules.

Figure 5.2 shows separate lines for each of the five possible strategies (always reject, preponderance of positive evidence, follow the single most reliable signal, preponderance of negative evidence, and always accept). As the figure shows, each

Fig. 5.2 When should a receiver follow two signals? This plot shows a modified version of Fig. 5.1, in which we have added expected payoff lines for the two "two-signal" rules. In general, the model predicts five different levels of p in which each of five possible strategies is best. The text gives the details.



of the five possible strategies gives the highest payoff in a different region of the p-axis. As p—the relative frequency of the good condition—increases from zero to one, we see an orderly progression in which each of the five possible rules is best in a different interval of p value (we call these intervals "layers" below). We can see that there exist four values of p, $p_1 < p_2 < p_3 < p_4$, such that the following is true:

Layer 1: $0 \leq p < p_1$: Always reject
Layer 2: $p_1 < p < p_2$: Use the preponderance of positive evidence rule
Layer 3: $p_2 < p < p_3$: Follow the single most reliable signal
Layer 4: $p_3 < p < p_4$: Use the preponderance of negative evidence rule
Layer 5: $p_4 < p \leq 1$: Always accept.

Notice that the preponderance of positive evidence rule can only pay off when $p < 1/2$. This occurs because when $p < 1/2$, it is, on average, best to reject. This makes sense because the preponderance of positive evidence rule is biased in favor of rejection. We can interpret the preponderance of positive of evidence rule as "normally reject, but accept if and only if two '+' signals are observed." Similarly, the preponderance of negative evidence rule can only pay off when $p > 1/2$ (when it is best, on average, to accept), and we can interpret the preponderance of negative evidence rule as: "normally accept, but reject if and only if two '−' signals are observed." Finally, notice that when $p = 1/2$, it can never pay to use a combination of signals. We predict instead that when the good and bad conditions are equally likely, the animal should always follow the single most reliable signal.

It may be surprising to some that the base rate p, the probability of the good condition, or environmental uncertainty more broadly, is *the* critical element in making signal combinations important. In our model, combinations can become important because they can tell you when to depart from the default (best on average) behavior. This is quite different from the folk wisdom about the statistical value of multiple signals. People tend to think that multiple sources of information will necessarily lead to better decisions, but this is not generally true. The value of signal combinations here derives from their power to indicate when you should deviate from your default action. In the absence of a default action (i.e., when $p = 1/2$) you should never use multiple signals but instead follow the single most reliable source of information. Moreover, there is a sense in which the two "two-signal rules" (preponderance of positive and negative evidence) represent a relatively weak form of signal use. In these rules, the animal only changes its behavior for certain specific and special signal combinations, otherwise it proceeds merrily along following its default "averaging" behavior. In contrast, an animal following the single most reliable signal matches its behavior to the signal every time the signal is observed.

5.4 Effects of Environmental Uncertainty and Signal Reliability

While the results above give us the general qualitative form of the solution—that is, the five layers—we would like to know how our focal variables [environmental uncertainty (p) and signal reliability (q and r)] influence these layers. To answer this question, we seek algebraic expressions for the four break points (p_1, p_2, p_3, p_4). This is not difficult mathematically, because each break point is defined by the intersection of two well-specified straight lines. Figure 5.3 uses these calculations to show the optimal strategies across a range of values for our three key variables: the reliability of the most reliable component S (given by q), the reliability of the less reliable component T, (given by r, $q > r$), and the relative frequency of the good condition (p). The three panels of Figure 5.3 show the plots of r versus p at each of three separate values of q (Fig. 5.3a, low $q = 0.65$; Fig. 5.3b, intermediate $q = 0.795$; and Fig. 5.3c, high $q = 0.99$). Notice that the scale of the r-axis varies between the plots even though each plot occupies the same visual space; r ranges from 0.5 to q, because the reliability of the second most reliable component is necessarily less than or equal to the reliability of the most reliable component. The plots show that our five regions are each roughly triangular and that they nest together in a tooth-like way (though these trends become somewhat distorted at high values of q). We make three observations about this figure. First, notice that when $r = 1/2$, so that the second signal is completely unreliable, our model is identical to the single-signal model. The animal should always accept if $p > q$, use the signal if $(1 - q) < p < q$, and always reject if $p < (1 - q)$ (just as in Fig. 5.1). Second, as the reliability of the second signal (r) increases, the layers in which signal combinations matter (layers 2 and 4) also increase in size, as we would expect. When the two components are equally reliable ($r = q$), it no longer makes sense to follow the most reliable component. As one might expect, the mathematics

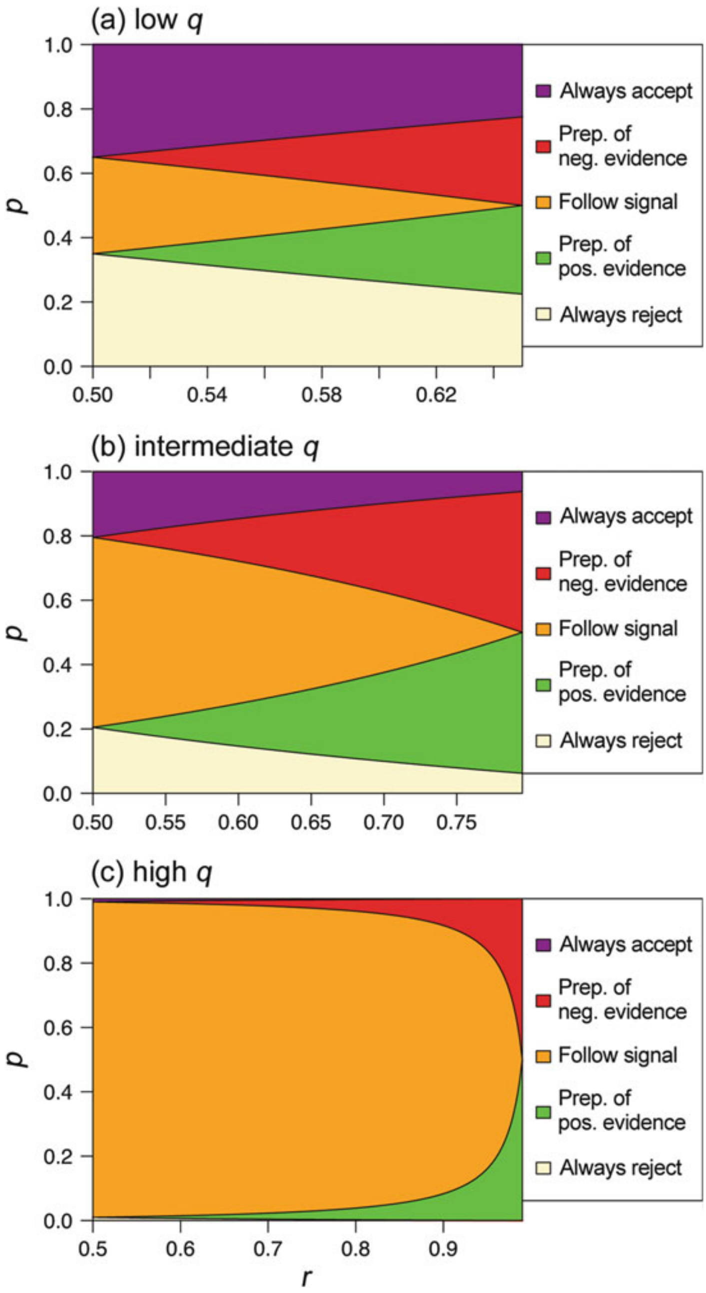


Fig. 5.3 The joint effects of environmental uncertainty and signal reliability on following two signals. Optimal strategies as a function of p and r, at three different levels of q: (**a**) $q = 0.65$, (**b**) $q = 0.795$, and (**c**) $q = 0.99$. Note that the scale of the x-axis changes between plots; since the reliability of the less reliable component (r) is lower than the reliability of the more reliable component (q), the x-axis ranges from $0.5 \leq r \leq \mathrm{q}$

reveals a limiting case in which the "follow the most reliable" region disappears. In this case, when signal following pays, it always pays to attend to both signal components. Specifically, whenever $p > 1/2$, the animal should either follow the preponderance of negative evidence rule or always accept, depending on the precise level of p; whenever $p < 1/2$, the animal should either follow the preponderance of positive evidence rule or always reject, again depending on the precise level of p. Finally, notice that as q approaches 1, the conditions that favor multiple signal use become increasingly narrow.

5.5 Environmental Uncertainty and Complex Signaling

With this model, we have illustrated that receivers can benefit economically from following two signals at certain combinations of environmental uncertainty (p) and signal reliability (q and r). Generally speaking, two-signal strategies are favored at intermediate levels of environmental uncertainty, that is, when the environment is neither highly certain ($p = 0$ or $p = 1$) nor highly uncertain ($p = 0.5$). This raises many questions about whether specific examples of complex signals are indeed correlated with intermediate levels of uncertainty as we hypothesize. While we suspect that new empirical studies will be needed to adequately test our hypothesis, we argue that there is broad qualitative support for our central predictions in the signaling literature.

Our first prediction is that signals should not be used at all in highly certain conditions ($p = 0$ and $p = 1$). Instead, animals should adopt inflexible strategies. In general, this appears to be true; we do not expect to see communication when a receiver should always do the same thing. For example, female ornaments are uncommon in traditional mating systems; typically, males maximize their fitness by mating with all available females (that is, they should follow the rule "always mate"), so females need not invest in costly courtship displays. Signaling should arise in uncertain conditions, when receivers can use signals to inform a decision between multiple behavioral responses. Specifically, our model predicts that intermediate uncertainty ($p > 0.5$ or $p < 0.5$) will favor complex signals while maximal uncertainty ($p = 0.5$) will favor simple signals. Quantifying uncertainty in natural systems is a nontrivial task; researchers must define "good" and "bad" states and determine their relative abundances. Some signaling systems will be more amenable to this classification scheme than others; ideally, states should be binary, clearly delineated, and easily categorized by researchers.

As an illustrative example of intermediate uncertainty, consider identity signals. Identity signals function to facilitate individual recognition, and should arise when it is favorable for a signaler to be correctly identified by conspecifics (Johnstone 1997; Tibbetts and Dale 2007; Chaps. 7 and 8). Upon encountering a conspecific, a receiver must determine if that individual is, say, "Frank" or "not Frank" based on an identity signal. (Note that this signal fits our criteria nicely—it is binary, clearly delineated, and easily observed by researchers.) Uncertainty will always be skewed in such signals; there is only one Frank in a population of size n, so the value of p will be $1/n$. Our model predicts, then, that identity signals should have multiple components and that is indeed what we observe. The need for multiple characteristics is intuitive in individual recognition; at a minimum, individuals must vary in enough qualities that they can be identified as unique. Researchers have shown that increasing population size results in increased signal variation (Pollard and Blumstein 2011, 2012) and have further argued that the need for individual recognition should promote or maintain signal diversity (Beecher 1989; Dale et al. 2001; Tibbetts 2004; Tibbetts and Dale 2007). Individual recognition is seen across a range of taxa, and different species rely upon different signal types; however, multicomponency of identity signals is ubiquitous. Such signals can range from complex chemical profiles (e.g., Steiger et al. 2008; delBarco-Trillo et al. 2012), to visual characteristics (e.g., Dale et al. 2001; Tibbetts 2002), to vocalizations (e.g., Clark et al. 2006; Chaps. 7 and 8), and, of course, traits across multiple modalities (e.g., Proops et al. 2008; Kondo et al. 2012).

Our model predicts that maximum uncertainty ($p = 0.5$) should favor simple signals. Perfect uncertainty (like any precise value) is likely to be rare in natural systems; however, some signal types might be more likely to occur near $p = 0.5$. For example, sex recognition signals might meet this criterion since sex ratios are often more or less balanced. Somewhat surprisingly, it is difficult to find evidence for “simple signals” in the existing literature. Many traditional communication studies focus on a single attribute of a signal; indeed, this is a popular criticism among complex signaling researchers (e.g., Hebets and Papaj 2005). However, such studies rarely seek to verify that receivers respond to *only* one component. Though complex signaling studies often attempt to isolate the effects of individual components, researchers may choose to study likely complex signals or treat simple signal following behavior as a negative result; either of these could result in a publication bias against reporting the existence of simple signals. There is, however, some support for our general prediction in lab-based, learned signal-following studies that test responses in perfectly uncertain environments (i.e., $p = 0.5$). Perfect uncertainty is a common experimental condition in learning studies because it is a theoretically important case and because perfect uncertainty maximally favors learning. Rubi and Stephens (2016) examined the responses of blue jays (*Cyanocitta cristata*) to two reliable signal components (a color and a pattern) at various combinations of q and r and found that receivers followed only one component (typically the more reliable one). Kazemi et al. (2014) found that blue tits (*Cyanistes caeruleus*) could learn to avoid color, pattern, and shape, but that color overshadowed the effect of the other two components in compound learning. In experiments on domestic chicks (*Gallus gallus domesticus*), Aronsson and Gamberale-Stille (2008) found that color overshadowed pattern, and Siddall and Marples (2011) found that color overshadowed auditory cues; however, in both of these studies the less salient stimulus was followed weakly or not at all, indicating that this behavior may have resulted from sensory constraints rather than informational strategies. More naturalistic studies necessarily have less control over parameters such as uncertainty and reliability; however, improved technology has made it possible to more thoroughly characterize what aspects of a signal are important for receivers. For example, Yorzinski et al. (2013) used telemetric gaze trackers to determine that peahens (*Pavo cristatus*) prioritize certain components of a peacock's courtship display and ignore other, highly conspicuous components. This example illustrates the danger of assuming that receivers will utilize all seemingly important components of a signal.

5.6 Economic Benefits Versus Psychological Benefits

In the complex signaling literature, sensory, perceptual, and cognitive benefits are often presented as alternatives to economic benefits. There is, however, no reason why these approaches should be considered mutually exclusive. It is perfectly plausible that complex signals confer both psychological and economic benefits to receivers simultaneously. While these may be truly independent effects, we argue that our approach may provide another explanation for why receiver brains seem to respond so strongly to complex signals. In psychology, multiple stimuli are traditionally discussed in the context of constraints. Classic learning theory predicts that multiple stimuli will interfere with each other by competing for associative strength (Rescorla and Wagner 1972) or receiver attention (Mackintosh 1975). Blocking and overshadowing are classic examples in which learning about one stimulus actively inhibits learning about a second, and studies focusing on divided attention often find deficits in performance when multiple components are present (Treisman and Gelade 1980; Dukas and Kamil 2000, 2001; Palmer et al. 2000; Clark and Dukas 2003). Thus, the classic view is that multicomponent stimuli present cognitive challenges, rather than cognitive advantages, to receiver processing. However, cognitive and perceptual processes are evolved traits like any other, and we expect natural and sexual selection to act on these processes when possible. When following multiple signal components is advantageous, selection should favor sensory systems and neural circuitry that minimize constraints or otherwise facilitate processing multiple stimuli. Researchers typically focus on the efficacy advantages of multicomponent signals such as improved signal reception; the economic advantages outlined here are another plausible driver of selection.

5.7 Limitations of this Approach

We acknowledge that our model makes some large simplifying assumptions that will need to be considered when applying it to specific natural systems. First, our model assumes dichotomous "good" and "bad" signaler states. This classification scheme will work for some signal types, such as identity signals (Sect. 5.5). Classification becomes trickier with qualities that vary more continuously, as is likely, for example, when signals are used to assess mate quality or fighting ability. Determining a threshold for a "good male" can be challenging. We have also assumed a highly simplified payoff structure in which any correct response is equally beneficial and any error is equally costly. The consequences of errors will vary greatly across systems; a missed mating opportunity is not the same as a missed foraging opportunity (Wiley 2015). Even within a system, the costs of different errors will likely differ; failing to consume a palatable prey item is likely less bad than consuming a poisonous one (see Chap. 11). Finally, this model assumes that receiver strategies are fixed across individuals and across time. There is good evidence, however, that receiver choosiness can vary with intrinsic factors, such as receiver need (Brower and Calvert 1985; Skelhorn and Rowe 2007).

5.8 Summary and Future Directions

While complex signals offer many perceptual and cognitive advantages over single component signals, economic models typically suggest that extra components provide no informational benefit to receivers. We offer a new model that contradicts this economic claim. We identify situations in which it pays to attend to signal combinations even when these combinations offer no psychological advantage. Our model suggests that it can pay to attend to signal combinations when each of two-signal components indicates that an animal should deviate from its normal or default behavior. This, in turn, depends on the underlying "base rate" of the conditions that animals signal about. For example, if males signal about their quality and poor quality males are more common, then receivers can benefit from responding to a combination of signals that jointly indicate the presence of a rare high-quality male. In contrast, if good and bad males were equally likely, we would expect receivers to follow the single most reliable signal component and ignore signal combinations. These predictions would be relatively straightforward to test in lab-based signaling games. In such systems, experimenters can manipulate the base rate of conditions. Once the receivers have adopted stable signal-following behavior, the experimenters can test responses to all combinations of components to characterize the strategy that the receivers adopt. Following the example we have laid out in this chapter, the base rate would be the proportion of the time the condition is "good." Quantifying the required response variables outlined in Table 5.2 (a, b, c, and d) would be a simple matter of testing responses to all four component combinations. These variables would then be used to characterize the adopted strategy.

We found that attending to combinations was beneficial when signaler states are asymmetrical, and we hypothesize that the required asymmetries exist in many natural signaling problems, such as mate quality signals, signals of fighting ability, and individual recognition signals. We do not offer this model as an alternative to the many documented perceptual and cognitive benefits of complex signals. Instead, we hypothesize that the economic advantages identified here will often work hand-in-hand with the well-documented psychological advantages of complex signals, giving us a new and more nuanced explanation of the prevalence of complex signals in nature.

Acknowledgements We thank Tim Polnaszek and Virginia Heinen for helpful discussions on the model and the editors and an anonymous reviewer for comments on the manuscript.

References

Aronsson M, Gamberale-Stille G (2008) Domestic chicks primarily attend to colour, not pattern, when learning an aposematic coloration. Anim Behav 75:417–423. doi:10.1016/j.anbehav.2007.05.006

Beecher MD (1989) Signalling systems for individual recognition: an information theory approach. Anim Behav 38:248–261. doi:10.1016/S0003-3472(89)80087-9

Bro-Jørgensen J (2010) Dynamics of multiple signalling systems: animal communication in a world in flux. Trends Ecol Evol 25:292–300. doi:10.1016/j.tree.2009.11.003

Brower LP, Calvert WH (1985) Foraging dynamics of bird predators on overwintering monarch butterflies in Mexico. Evolution 39:852–868

Candolin U (2003) The use of multiple cues in mate choice. Biol Rev Camb Philos Soc 78:575–595

de Caprona MC, Ryan M (1990) Conspecific mate recognition in swordtails, *Xiphophorus nigrensis* and *X. pygmaeus* (Poeciliidae): olfactory and visual cues. Anim Behav 39:290–296

Clark C, Dukas R (2003) The behavioral ecology of a cognitive constraint: limited attention. Behav Ecol 14:151–156

Clark JA, Boersma PD, Olmsted DM (2006) Name that tune: call discrimination and individual recognition in Magellanic penguins. Anim Behav 72:1141–1148. doi:10.1016/j.anbehav.2006.04.002

Dale J, Lank D, Reeve H (2001) Signaling individual identity versus quality: a model and case studies with ruffs, queleas, and house finches. Am Nat 158:75–86

Deag J, Scott G (1999) "Conventional" signals in avian agonistic displays: integrating theory, data and different levels of analysis. J Theor Biol 196:155–162. doi:10.1006/jtbi.1998.0825

delBarco-Trillo J, Sacha CR, Dubay GR, Drea CM (2012) Eulemur, me lemur: the evolution of scent-signal complexity in a primate clade. Philos Trans R Soc B Biol Sci 367:1909–1922. doi:10.1098/rstb.2011.0225

Dukas R, Kamil AC (2000) The cost of limited attention in blue jays. Behav Ecol 11:502–506. doi:10.1093/beheco/11.5.502

Dukas R, Kamil AC (2001) Limited attention: the constraint underlying search image. Behav Ecol 12:192–199. doi:10.1093/beheco/12.2.192

Dunlap AS, Stephens DW (2014) Experimental evolution of prepared learning. Proc Natl Acad Sci 111:11750–11755. doi:10.1073/pnas.1404176111

Dunlap AS, Stephens DW (2009) Components of change in the evolution of learning and unlearned preference. Proc Biol Sci 276:3201–3208. doi:10.1098/rspb.2009.0602

Grether GF, Kolluru GR, Nersissian K (2004) Individual colour patches as multicomponent signals. Biol Rev Camb Philos Soc 79:583–610

Hebets EA, Papaj DR (2005) Complex signal function: developing a framework of testable hypotheses. Behav Ecol Sociobiol 57:197–214. doi:10.1007/s00265-004-0865-7

Johnstone RA (1995) Honest advertisement of multiple qualities using multiple signals. J Theor Biol 177:87–94. doi:10.1016/S0022-5193(05)80006-2

Johnstone RA (1997) Recognition and the evolution of distinctive signatures: when does it pay to reveal identity? Proc R Soc B Biol Sci 264:1547–1553. doi:10.1098/rspb.1997.0215

Kazemi B, Gamberale-Stille G, Tullberg BS, Leimar O (2014) Stimulus salience as an explanation for imperfect mimicry. Curr Biol 24:965–969. doi:10.1016/j.cub.2014.02.061

Kim S-Y, Noguera JC, Morales J, Velando A (2011) The evolution of multicomponent begging display in gull chicks: sibling competition and genetic variability. Anim Behav 82:113–118. doi:10.1016/j.anbehav.2011.04.005

Kondo N, Izawa E-I, Watanabe S (2012) Crows cross-modally recognize group members but not non-group members. Proc R Soc B Biol Sci 279:1937–1942. doi:10.1098/rspb.2011.2419

Leonard M, Horn A, Parks E (2003) The role of posturing and calling in the begging display of nestling birds. Behav Ecol Sociobiol 54:188–193. doi:10.1007/s00265-003-0626-z

Mackintosh NJ (1975) A theory of attention: Variations in the associability of stimuli with reinforcement. Psychol Rev 82:276–298. doi:10.1037/h0076778

Massaro DW, Cohen MM (1990) Perception of synthesized audible and visible speech. Psychol Sci 1:55–63. doi:10.1111/j.1467-9280.1990.tb00068.x

McLinn CM, Stephens DW (2006) What makes information valuable: signal reliability and environmental uncertainty. Anim Behav 71:1119–1129. doi:10.1016/j.anbehav.2005.09.006

McLinn CM, Stephens DW (2010) An experimental analysis of receiver economics: cost, reliability and uncertainty interact to determine a signal's value. Oikos 119:254–263. doi:10.1111/j.1600-0706.2009.17756.x

Page RB, Jaeger RG (2004) Multimodal signals, imperfect information, and identification of sex in red-backed salamanders (Plethodon cinereus). Behav Ecol Sociobiol 56:132–139. doi:10.1007/s00265-004-0774-9

Palmer J, Verghese P, Pavel M (2000) The psychophysics of visual search. Vision Res 40:1227–1268

Partan SR, Marler P (1999) Communication goes multimodal. Science 283(5406):1272–1273

Pearson DL (1989) What is the adaptive significance of multicomponent defensive repertoires? Oikos 54:251–253

Pollard KA, Blumstein DT (2011) Social group size predicts the evolution of individuality. Curr Biol 21:413–417. doi:10.1016/j.cub.2011.01.051

Pollard KA, Blumstein DT (2012) Evolving communicative complexity: insights from rodents and beyond. Philos Trans R Soc Lond B Biol Sci 367:1869–1878. doi:10.1098/rstb.2011.0221

Proops L, McComb K, Reby D (2008) Cross-modal individual recognition in domestic horses (*Equus caballus*). Proc Natl Acad Sci 106:1–5

Rescorla RA, Wagner ARA (1972) A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: Black AH, Prokasy WF (eds) Classical conditioning II: current research and theory. Appleton Century Crofts, New York, NY, pp. 64–99

Rowe C (1999) Receiver psychology and the evolution of multicomponent signals. Anim Behav 58:921–931. doi:10.1006/anbe.1999.1242

Rowe C, Halpin C (2013) Why are warning displays multimodal? Behav Ecol Sociobiol 67:1425–1439. doi:10.1007/s00265-013-1515-8

Rowe C, Skelhorn J (2004) Avian psychology and communication. Proc R Soc B Biol Sci 271:1435–1442. doi:10.1098/rspb.2004.2753

Rubi TL, Stephens DW (2016) Should receivers follow multiple signal components? An economic perspective. Behav Ecol 27:6–44. doi:10.1093/beheco/arv121

Schluter D, Price T (1993) Honesty, perception and population divergence in sexually selected traits. Proc R Soc B Biol Sci 253:117–122. doi:10.1098/rspb.1993.0089

Siddall EC, Marples NM (2011) Hear no evil: The effect of auditory warning signals on avian innate avoidance, learned avoidance and memory. Curr Zool 57:197–207

Skelhorn J, Rowe C (2007) Predators' toxin burdens influence their strategic decisions to eat toxic prey. Curr Biol 17:1479–1483. doi:10.1016/j.cub.2007.07.064

Steiger S, Franz R, Eggert A-K, Müller JK (2008) The Coolidge effect, individual recognition and selection for distinctive cuticular signatures in a burying beetle. Proc Biol Sci 275:1831–1838. doi:10.1098/rspb.2008.0375

Tibbetts EA (2004) Complex social behaviour can select for variability in visual features: a case study in Polistes wasps. Proc R Soc B Biol Sci 271:1955–1960

Tibbetts EA (2002) Visual signals of individual identity in the wasp *Polistes fuscatus*. Proc R Soc B Biol Sci 269:1423–1428. doi:10.1098/rspb.2002.2031

Tibbetts EA, Dale J (2007) Individual recognition: it is good to be different. Trends Ecol Evol 22:529–537. doi:10.1016/j.tree.2007.09.001

Treisman AM, Gelade G (1980) A feature-integration theory of attention. Cogn Psychol 12:97–136. doi:10.1016/0010-0285(80)90005-5

Wiley RH (2015) Noise matters: the evolution of communication. Harvard University Press, Cambridge, MA

Wilson AJ, Dean M, Higham JP (2013) A game theoretic approach to multimodal communication. Behav Ecol Sociobiol 67:1399–1415. doi:10.1007/s00265-013-1589-3

Yorzinski JL, Patricelli GL, Babcock JS et al (2013) Through their eyes: selective attention in peahens during courtship. J Exp Biol 216:3035–3046. doi:10.1242/jeb.087338

Chapter 6
Communication Through a Window of Error: Proportional Processing and Signal Categorization

Karin L. Akre and Sönke Johnsen

Abstract Animals can perceive information about the physical and social environment around them only through their sensory systems. Because of this, the constraints of sensory mechanisms can impact animals' abilities to accurately assess their environment. In communication systems, receivers depend on perceiving signals to make decisions about how to respond. One feature of sensory systems that can blur their perceptual interpretation of stimuli is an imperfect ability to compare magnitudes. This complicates communication tasks such as signal categorization. For many sensory parameters, this difficulty is exacerbated for high-magnitude stimuli. Often, the potential error in perceiving stimulus magnitude varies predictably according to magnitude. This allows us to predict how receiver error can influence the evolution of communication systems, including signal characteristics, signaler strategies, and receiver behaviors. Perceptual error can contribute to the evolution of exaggerated signals, novel signal components, loss of signal components, and directional changes according to signal function. For signalers, perceptual error by receivers can lead to changes in rate of display, choice of dynamic signal components, and escalation of competitions. For receivers, this process sheds light on the evolution of sensory systems, allocation of attention, and compensatory behaviors. We invite researchers to apply these concepts in diverse areas of animal communication study.

K.L. Akre (✉)
Psychology Department, Hunter College, New York, NY 10065, USA
e-mail: karin.akre@gmail.com

S. Johnsen
Biology Department, Duke University, Durham, NC 27708, USA
e-mail: sjohnsen@duke.edu

M.A. Bee, C.T. Miller (eds.), *Psychological Mechanisms in Animal Communication*, Animal Signals and Communication 5,
DOI 10.1007/978-3-319-48690-1_6

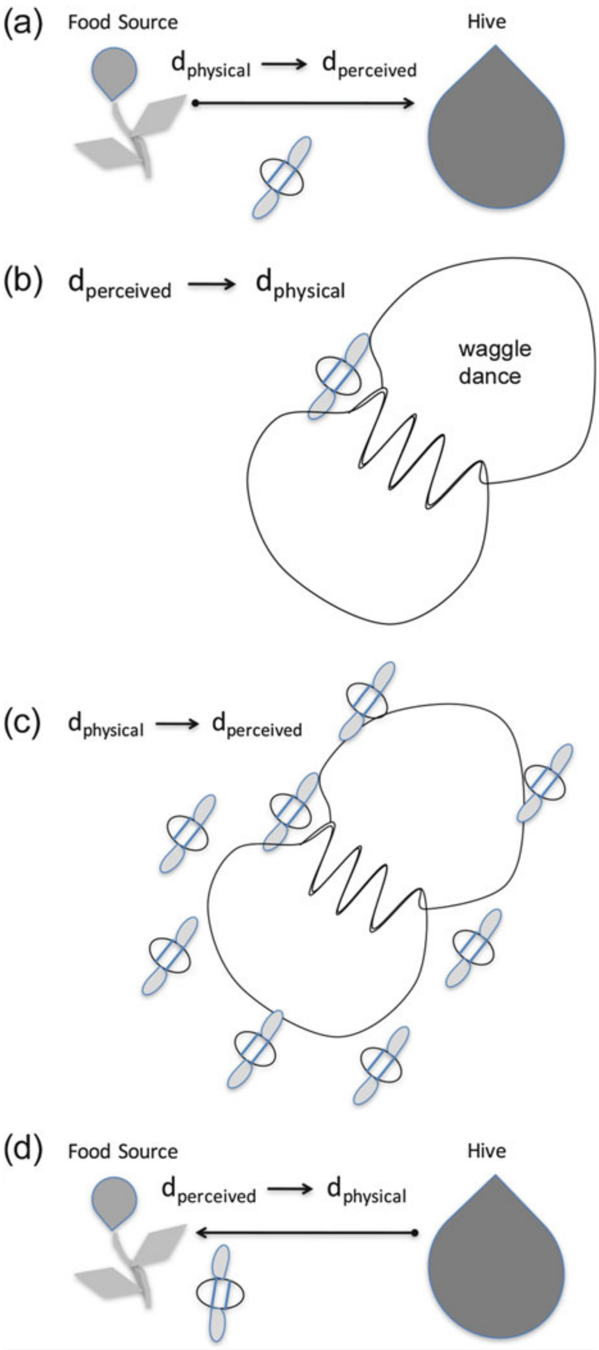
6.1 Sensory Perception and Just-Noticeable Differences

The potential for error in translating between physical stimulus values and perceived ones can impact the evolution of animal communication (Akre and Johnsen 2014). When animals communicate, they produce signals that influence the behavior of others. This process depends upon receiver sensory perception of the signals; thus, animals have evolved an incredible diversity of mechanisms to detect and process sensory parameters ranging from the polarization properties of light (Cronin et al. 2014) to the temporal features of electric pulses (Bullock et al. 2006). In every mechanism for every sensory parameter, there is a window of error in identifying the magnitude of the signal features and in determining whether two signals differ from each other. The *just-noticeable difference* (*JND*) is the minimum difference in stimulus magnitude that can be detected. Because sensory systems cannot avoid JNDs, receivers can mismeasure communication signals. This means that they can categorize a stimulus incorrectly and thus respond incorrectly. Since mistaken responses can impact both signaler and receiver fitness, evolution is likely to respond to this situation. Decades of research show that receiver nervous systems evolve in response to the features of signals, and the evolution of signal features can be shaped by receiver perception and response (Guilford and Dawkins 1991; Miller and Bee 2012; Rowe 2013; Chap. 2). The evolutionary impact of JNDs should also be substantial. Signalers can evolve to make their signals easier to identify. Signals can evolve to be more easily perceived. Receivers can evolve to get better at discriminating between signals. But still, there will always be JNDs.

Our broad aim in this chapter is to examine how evolution responds to the mismeasure of sensory information and how understanding the nature of receivers' JNDs can help us to predict the evolution of animal communication systems. We describe how the errors made by receivers often follow consistent patterns and how this allows us to recognize and predict evolutionary patterns. We then examine how receiver error is predicted to impact the evolution of signals, focusing on halting the evolution of signal magnitude increase, adding new signal components, losing original signal components, and moving signals directionally over time. We then examine how receiver error could impact the evolution of signaler behaviors, focusing on behavior parallels to the evolutionary trajectories of signals, and how signalers should balance trade-offs between high-magnitude and mid-level signals maintained over time. We then consider how receiver error could impact the evolution of receiver behaviors, considering the distribution of receiver attention, the possibility of compensatory behaviors, and the impact of these ideas on signal detection theory. Finally, we describe variability within patterns of JNDs and discuss some future directions for researchers who want to pursue these ideas.

To illustrate how JNDs can impact animal communication systems, consider the foraging honeybee (*Apis mellifera*). A honeybee can communicate the direction and distance to food sources by waggle dancing in a figure 8 pattern inside its hive (Fig. 6.1). Attentive foragers in the hive perceive the dance and can find the described food when they leave the hive (von Frisch 1967). In this process, a

Fig. 6.1 An illustrative example of the problem of potential errors resulting from proportional processing. In the dance communication of honeybees (*Apis mellifera*), individuals rely on distance and direction information that is transformed multiple times between physical and perceived values. Here we show the many transformations of distance (d) information in this process. (**a**) First, a forager flies a physical distance and remembers a perceived distance. (**b**) Then that perceived distance is communicated in a physical dance. (**c**) Then, the physical dance duration is perceived by attending foragers. (**d**) Finally, those foragers physically fly the distance they perceived. At each step, some error is introduced from the translation between physical and perceived distance



transformation between physical and perceived distance and direction occurs repeatedly. The dancer perceives actual physical distance and direction when traveling home from the food (Fig. 6.1a). She then translates the perceived values to physical signal properties in the dance (Fig. 6.1b), which are perceived by attentive foragers (Fig. 6.1c). These foragers then travel an actual physical distance and direction to reach the food (Fig. 6.1d). The foragers rely on careful translation between physical and perceived stimulus properties, as all animals must in order to interact with their environment. Yet the transformation from physical to perceived stimulus magnitude is never entirely accurate. Every physical sensory parameter is measured perceptually within a window of error discussed above, the JND. So, the receiver foragers' perception of duration, length, and orientation of a honeybee waggle dance will never be precisely what is physically produced. Any error in perception can cause animals to make mistakes. For example, foraging honeybees sometimes fall short of or overshoot the correct distance to a flower (Riley et al. 2005).

One problem receivers must solve when responding to the perceived values of physical signals is identifying the ranges of signal variation that correspond with behaviorally relevant categories. For example, a female frog at a mixed species chorus must distinguish between the vocalizations of potential conspecific mates and heterospecific individuals (Ryan and Rand 1993; Gerhardt 1994). A territorial male frog must determine whether the vocalizations it hears are from a neighbor or a stranger in order to avoid wasting energy on attacking a neighbor (Bee and Gerhardt 2002; Chap. 7). Distinguishing between categories like these is easier when sensory systems have smaller JNDs. Depending on how animals process signals, the relevant sensory parameter of the vocalizations in question could be frequency, duration, amplitude, or some other acoustic feature (Beecher et al. 1979; Charrier et al. 2002). With a small JND, the perceived value of the relevant sensory parameter will be close to the actual stimulus value. When JNDs are large, the perceived value could be farther from the actual stimulus value. Because of this, larger JNDs can make it difficult to notice the difference between two dissimilar stimuli.

When receivers have large JNDs, they might categorize communication signals incorrectly. For example, a female frog that is listening to vocalizations in order to identify an appropriate mate might hear a continuous spectrum of variation in call frequency at a chorus. Hearing the difference between heterospecific and conspecific calls might be difficult if their calls are similar. Together, the size of her JND and the difference between the two species calls will determine whether she reliably chooses conspecifics as mates. If choosing a heterospecific mate is costly, the two species could evolve reproductive character displacement by which the signals evolve to become more distinct (Waage 1975; Servedio and Noor 2003). The species could also achieve reproductive isolation by receivers evolving to be more selective through changes in how they distinguish between the two call types (Gerhardt 1994; Gabor and Ryan 2001). Understanding JNDs should help us understand the evolution of signals, signal assessment, and how accurately animals categorize behaviorally relevant signals.

6.2 Proportional Processing and Receiver Error

When measuring the difference in magnitude between two stimuli, one can consider both absolute differences and proportional differences. Actually counting the berries in two baskets of fresh strawberries could reveal that one has 25 strawberries and the other has 20. The absolute difference is a subtraction problem—25 minus 20 equals 5 strawberries. The proportional difference is a matter of division—an absolute difference of five strawberries divided by the total 25 strawberries in the larger basket equals 1/5 of the strawberries. Both ways of assessing difference allow us to distinguish between the two baskets and know which one is larger.

For most sensory parameters, the nervous system assesses proportional differences to distinguish between the magnitudes of two stimuli (Akre and Johnsen 2014). For example, visually determining which of two lights is brighter depends on comparing their proportional difference (Fig. 6.2). Suppose that a building has

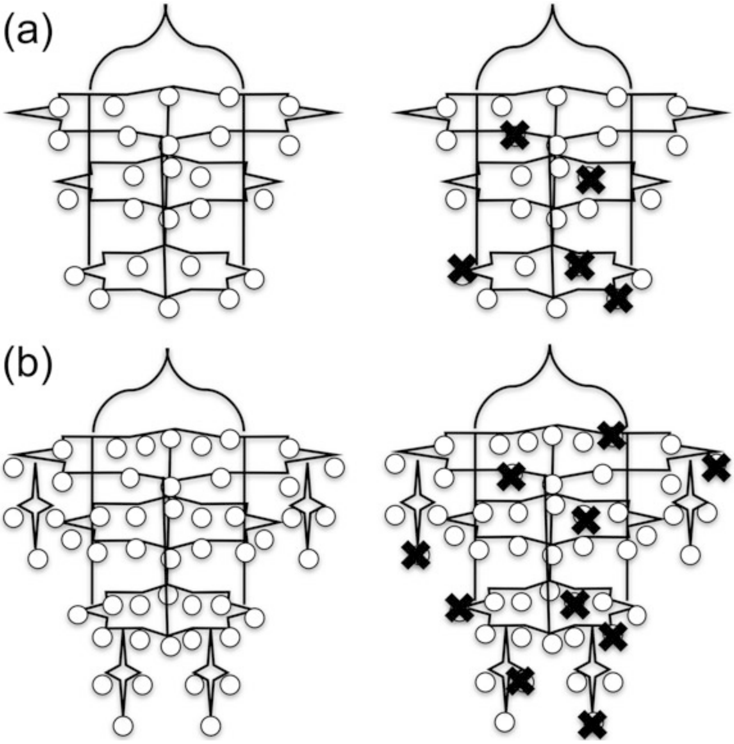


Fig. 6.2 An illustrative example of the Weber fraction and just-noticeable difference (JND). According to Weber's law, brightness has a JND (ΔI) that remains a constant proportion (k) of the total stimulus magnitude (I) for much of the variation we can perceive (Cronin et al. 2014). Here, the Weber fraction (k) is 1/5 of the total stimulus magnitude (I). Thus, the absolute value of the JND (ΔI) is smaller for lower-magnitude stimuli. In (**a**) 1/5 of 25 bulbs is 5 bulbs, and in (**b**) 1/5 of 50 bulbs is 10 bulbs

chandeliers that use 25 bulbs to light the rooms. If one bulb burns out, nobody notices, but when five bulbs burn out (1/5 of 25), people notice the room is dimmer than other rooms (Fig. 6.2a). This means there is a JND of about 1/5 the total brightness magnitude. Since the JND is defined by a proportional difference from the original magnitude, the absolute size of the JND for brightness varies systematically with stimulus magnitude. For many sensory parameters, the absolute size of the JND is a constant proportion of the stimulus magnitude for at least part of its range. So, in another building with more ornate chandeliers that use 50 bulbs, 10 bulbs (again 1/5 the total) would have to burn out before anyone noticed that the room was dimmer than the other rooms (Fig. 6.2b). When the JND is a constant proportion of the original stimulus magnitude, the absolute size of the JND is necessarily smaller at low magnitudes, while at high magnitudes, the absolute size of the JND is larger. The same proportional difference—1/5 in the case of our chandeliers—is an absolute difference of 5 bulbs in the first case and 10 bulbs in the latter (Fig. 6.2). Sensory systems with JNDs that vary like this are said to follow Weber's law (Weber 1834; Fechner 1860):

$$\frac{\Delta I}{I} = k, \text{or alternatively}, \Delta I = kI,$$

where ΔI is the JND in absolute terms of a difference in stimulus magnitude and I is the original stimulus magnitude. The constant k, known as the Weber fraction, indicates that the absolute JND is a constant proportion of the original stimulus magnitude. In the two examples illustrated in Fig. 6.2, ΔI and I differ, but the Weber fraction ($k = 1/5$) remains constant. We refer to perception that follows Weber's law as proportional processing.

When signal magnitude helps animals decide how to respond appropriately, proportional processing can cause receivers to err in predictable ways. Receiver mistakes are predictable because we know which types of signals are harder to distinguish—those produced at higher magnitudes by signalers or associated with larger absolute JNDs of receivers. Proportional processing influences how receivers perceive the stimulus variation relevant to categorization, biasing them to notice more variation in lower ranges of magnitude. This feature of signal categorization could be partially responsible for several nonlinear patterns of behavior, in which incremental changes in response to stimuli do not map linearly onto incremental changes in the stimuli themselves. For example, some females responding to differences in male frog calls, as in gray treefrogs, show nonlinear preferences for particular call characteristics: they are less discriminating of signal differences at larger absolute signal magnitudes (Gerhardt et al. 2000; Schwartz et al. 2001; Bush et al. 2002; Bee 2008). Generally, the likelihood of receivers mis-categorizing signals varies directly with stimulus magnitude.

Although this phenomenon makes errors predictable, it is important to point out that even for signal properties for which animals use proportional processing, Weber's law holds true only for certain ranges of magnitude. For example, humans show a "near miss" to Weber's law when discriminating high sound intensity (Yost 1994). Insects and anurans show an "opposite miss" to Weber's law under these circumstances (Gerhardt 1987; Forrest 1994; Wyttenbach and Farris 2004; Bee et al. 2012). At dim light levels, many vertebrates operate under a different law that still relies on proportional light comparisons (Cronin et al. 2014). For the current discussion, we refer specifically to ranges of sensory parameters that follow Weber's law. We discuss the predictability of other patterns' impacts on behavioral evolution in Sect. 6.6.

Proportional processing is a key aspect of receiver psychology that can influence the evolution of signals, signaling, and signal reception. When receivers follow Weber's law, higher-magnitude signals are harder to recognize, discriminate, and match. As a result, signals can evolve to influence receiver ability to perceive signal magnitude accurately. When signalers benefit from receiver mistakes, signalers and their signals might evolve to exploit receivers and encourage mistakes. On the other hand, when signalers benefit from receiver accuracy, both signals and signaler strategies can evolve to reduce receiver mistakes. Whether such patterns of evolution actually occur will depend not only on receiver perception but also on signal production costs. If costs increase linearly or exponentially with signal magnitude, the likelihood of adjusting signal magnitude due to receiver perception-based benefits will be quite different than when costs increase proportionally, or even decrease as signal magnitude increases. Also, receivers themselves should sometimes evolve to reduce mistakes in responding appropriately to signal variation. Receivers could do this by evolving reduced JNDs, adjusted attention, or compensatory behaviors (see Sect. 6.5).

In the remainder of this chapter, we discuss proportional processing-induced evolutionary changes in communication systems. Our goal is to explain how these changes are predicted to occur and include examples when possible. With this in mind, we will frame the ways proportional processing is expected to impact the evolution of animal communication and how research can address this topic. In Sect. 6.3, we describe how proportional processing can influence the evolution of physical features of signals. We show how it can halt sexual selection for increasing signal magnitude, trigger the evolution of new signal components, cause the loss of older signal components, and cause signals to evolve directionally. Next, in Sect. 6.4, we describe how signalers with dynamic signal displays can evolve strategies to optimize their influence on receivers that use proportional processing. In Sect. 6.5, we turn our attention to receivers and discuss how evolutionary changes in receiver behavior could reduce the probability of errors resulting from proportional processing. Section 6.6 briefly discusses exceptions in which transformations between actual and perceived signal magnitudes do not have proportional JNDs that are consistent across the full range of signal magnitude. Finally, in Sect. 6.7, we outline some directions for future research on proportional processing by receivers and animal communication systems.

In concluding this section, we note that few studies have yet applied the principles discussed in this chapter and hope that readers will look for potential applications in their own research. One significant challenge in elucidating when JNDs impact behavior is the occurrence of *just-meaningful differences* or *JMDs* (Nelson and Marler 1990). A JMD is a measure of what difference in signal magnitude is perceived by receivers as behaviorally significant (see also Chaps. 7 and 8). JMDs can depend upon an animal's motivational state, which can vary with hunger, reproductive cycle, age, or health condition. If a receiver does not respond to a difference in signal magnitude, it becomes necessary to consider whether the lack of response is due to not detecting the difference (due to its JND) or not caring about the difference (due to its JMD). Overcoming the difficulty of distinguishing between JNDs and JMDs is one of the challenges to tackle in proportional processing research, and we discuss potential solutions in Sect. 6.7.

6.3 Impacts on Signals

6.3.1 Halting the Evolution of Increasing Signal Magnitude

A well-established pattern in signal evolution is that females prefer more elaborate signals when making mate choices, leading to the evolution of elaborate male sexual signals (Andersson 1994; Jennions and Petrie 1997; Ryan 1998; Andersson and Simmons 2006). The term *elaborate* here refers to when females show preference for signals that are extended directionally in magnitude somehow—such as in intensity or in the number of components (Ryan 1985; Rowe 1999; Endler et al. 2005). An interesting twist to the evolutionary path of increasing elaboration in male signals is that proportional processing predicts that females are less likely to notice the same absolute differences in magnitude as male signals evolve to become higher magnitude (Cohen 1984). Several studies show nonlinear female preference for male traits, and frequently in these situations female preference is weaker at higher magnitudes. For example, gray treefrogs (*Hyla versicolor* and *Hyla chrysoscelis*) have a stabilizing preference for pulse rate and a directional preference for greater pulse number (Gerhardt et al. 2000; Schwartz et al. 2001; Bee 2008). Females discriminate the same absolute difference in pulse number less reliably at higher magnitude in both *H. versicolor* (Gerhardt et al. 2000) and *H. chrysoscelis* (Bee 2008). In *H. versicolor*, a small two-pulse difference that impacts behavior at low magnitudes has no effect when females discriminate between higher-magnitude stimuli (Schwartz et al. 2001). And in females of *H. chrysoscelis*, which exhibit a stabilizing preference for an average pulse rate, the strength of female preference for the average versus a lower alternative can only be matched in strength for comparisons with a higher alternative when the difference from the average to the higher call is much greater (Bush et al. 2002). In all these cases, there is a nonlinear preference function that might result from how females compare the calls they hear.

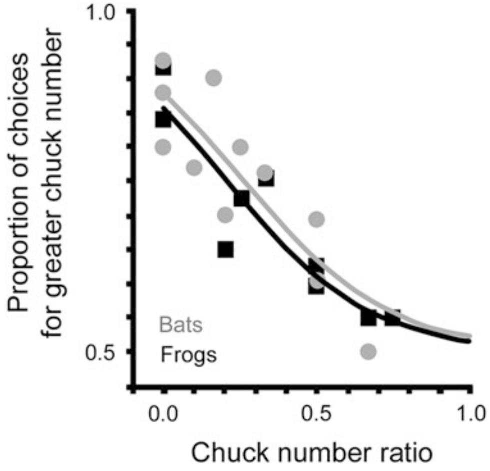


Fig. 6.3 Proportional processing of male túngara frog (*Physalaemus pustulosus*) calls by potential mates and predators. Both the preferences of female túngara frogs for a mate (*black circles* and *line*) and predatory attacks by fringe-lipped bats (*Trachops cirrhosus*; *gray circles* and *line*) target males that produce more chucks, significantly fitting a model based on proportional comparison of chuck number. Subjects were tested in nine different two-stimulus choice tests in which the paired call alternatives differed in absolute and relative chuck number. For female frogs, the ratios were 0:1, 0:3, 1:2, 1:3, 1:4, 1:5, 2:3, 2:4, and 3:4. The ratios for bats were 0:1, 0:2, 0:3, 1:2, 1:3, 1:4, 1:5, 1:6, 1:10, 2:3, and 2:4. (From Akre et al. 2011, permission granted by AAAS)

In some cases, female preference patterns clearly reflect Weber's law. In these cases, the female preference strength depends on the proportional difference between two call values. This specific situation has been demonstrated in túngara frogs, *Physalaemus pustulosus* (Akre et al. 2011). In túngara frogs, females prefer to mate with males that produce an extra element appended to the vocalization—a chuck (Chap. 4). Males can add 1–7 chucks but usually produce only a few chucks at natural choruses (Bernal et al. 2009). Males add chucks in response to female elicitation behaviors (Akre and Ryan 2011), but females do not consistently prefer more chucks with statistical significance (Bernal et al. 2009). Instead, they demonstrate a preference for more chucks that is constrained by their ability to discriminate chuck number due to their dependence on proportional processing (Fig. 6.3; Akre et al. 2011). Depending on the costs involved, this reduced benefit of adding chucks is likely to cause the evolution of increasing signal magnitude to cease.

The cost of signal production determines whether proportional processing will limit the evolution of increasing signal magnitude. When the cost of increasing signal magnitude is linear or exponential, such that each increment added requires an equal or increasing effort, the cost to benefit ratio at each additional increment becomes increasingly dominated by cost, because the benefit shrinks as females are less likely to notice the additional increment. When this occurs, proportional processing by females could be the reason that evolution of increasing signal magnitude ceases. The energetic requirement of producing additional signal increments is one type of cost that might increase in this way. But not all costs will increase with magnitude in this fashion. The fundamental frequency of some animals' vocalizations, for example, will increase in magnitude as body size decreases, such that signalers experience reduced cost with increasing magnitude. This would reduce the likelihood of limiting signal elaboration, even when the additional benefits for each increment are minimal.

Interestingly, predation pressure as a cost of signal production might not increase in a linear or exponential manner with additional signal increments. This is because changes in predation pressure, like the preferences of female túngara frogs, depend on perception of increasing physical stimulus magnitude. Predators might also experience difficulty in discriminating higher-magnitude signals from potential prey. When this is the case, predation pressure is not expected to increase dramatically with signal elaboration. In fact, this appears to be the case in túngara frog call evolution. Their major predator, the fringe-lipped bat (*Trachops cirrhosus*), discriminates between potential prey vocalizations by attending to chuck number, with a preference to attack individuals that produce more chucks (Chap. 11). The bats, like the female frogs, are constrained by proportional processing in their ability to actually attack the male making more chucks (Akre et al. 2011; see Fig. 6.3).

Generally, it is expected that predators drawn to conspicuous elaborate signals will limit the evolution of increasing signal elaboration (Zuk and Kolluru 1998). Predators might preferentially attack individuals that produce more elaborate signals due to factors such as ease of localization (Page and Ryan 2008) or quality of meal (Bernal et al. 2007). However, the impact of predation depends in part on their ability to discriminate signal magnitude. If predators face the same proportional processing constraints as conspecifics do, the impact of predation on signal evolution should actually decrease at higher magnitudes, as occurs with the predatory bats that target túngara frogs (Akre et al. 2011). In cases where eavesdropping predators have *smaller* JNDs than conspecific receivers, however, predators will be better than conspecific receivers at finding the higher-magnitude signals. This might be a common situation, because predators are usually bigger than prey and thus have bigger sense organs, which can lead to lower JNDs, as does the size of the eye in vision (Cronin et al. 2014). In this situation, the selective pressure posed by predation would increase, and predation would quickly halt the evolution of increasing signal elaboration. If predators instead have *larger* JNDs than conspecific receivers, predators will be relatively less capable of distinguishing between high-magnitude signals. In this situation, the selective pressure posed by predators would be negligible, and the evolution of increasing signal elaboration would depend completely on conspecific response (Fig. 6.4).

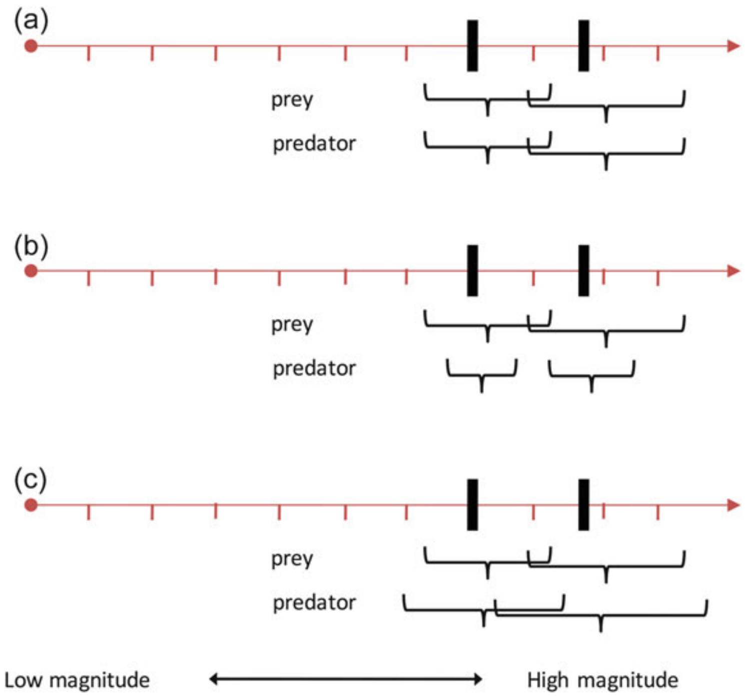


Fig. 6.4 The interaction between predator and prey discrimination thresholds and the evolution of increasing prey signal elaboration. In each image, the *horizontal axis* represents variation in stimulus magnitude along some sensory parameter, with equal linear increments marked. The *thick vertical marks* represent two stimuli having different magnitudes. The *horizontal brackets* beneath each stimulus represent the amount of change required to notice a difference from a particular stimulus value for prey (*above*) and predator (*below*) or the JND. When the JNDs of two signals overlap, receivers are sometimes unable to discriminate the two signals. Here we consider the evolutionary consequences when both conspecifics and predators preferentially seek out more elaborate signals. (**a**) When predators and prey have equal JNDs, both will experience the same proportional processing error, and both conspecific benefits and predation costs resulting from elaboration are reduced. The result is less elaboration. (**b**) When predators have smaller JNDs, predators will prefer elaborate signals, imposing high costs to elaboration, and signal elaboration will cease due to predation costs. (**c**) When predators have larger JNDs, predation will impose little cost to increasing increments, and signal elaboration will only cease due to limited benefits

6.3.2 New Signal Components

When the increasing magnitude of a signal component makes it difficult for receivers to distinguish between two signals, a decision could become costly for receivers. If choosing the higher-magnitude signal actually benefits a female, she will sometimes miss out on the benefit of making a correct choice. At the same time, the signalers suffer reduced benefits from producing high-magnitude signals because females cannot respond to additional signal increments. In a situation such as this, receivers could evolve a smaller discrimination threshold to better

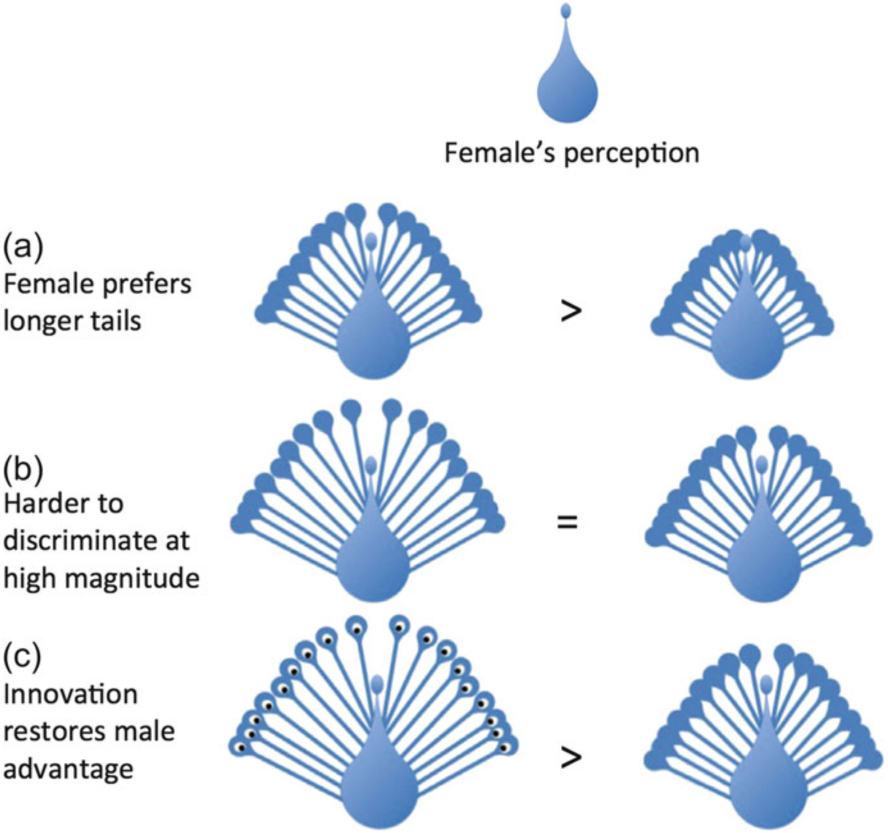


Fig. 6.5 Proportional processing and elaboration versus innovation in signals. (**a**) In a bird species in which females choose a male based on tail length, sexual selection could favor the elaboration of tails by making them longer. (**b**) However, proportional processing by females might make it difficult for females to distinguish between high-quality males with long tails. (**c**) The evolution of a novel, innovative trait, such as conspicuous spots on the tail, could allow high-quality males to recover an advantage over their competition. (From Akre and Johnsen 2014, permission granted by Elsevier)

perceive differences in signal magnitude. Another solution to this problem of ineffective communication is evolutionary change in the physical properties of the signal. If a signal no longer benefits signalers, signalers could evolve an entirely new signal component to be assessed along a separate axis that is not a continuous extension of previous complexity. Endler et al. (2005) describe the evolution of signals in terms of elaboration (extending the magnitude of a signal directionally) and innovation (changing signal magnitude in a new direction). The addition of an entirely novel signaling element can be thought of as innovation rather than elaboration (Fig. 6.5). This new signal component could occur in a separate sensory modality—adding a visual signal to an acoustic one (Chaps. 4 and 5)—or within a sensory modality by producing a categorically different element, as in bird songs with multiple elements strung together.

6.3.3 *Loss of Original Signal Components*

As described above, a species might evolve a novel signal component through innovation after the original component evolves to such high magnitude that it is difficult for receivers to discriminate between signals. In such cases, the benefit of producing the original signal component becomes questionable. Maintaining the original component as part of the signal might be critically important if receiver nervous systems require its presence as a precondition for recognizing and responding to the signal. Or, receivers might evolve to attend only to the new component, making the original component unnecessary. If production of the original component is at all costly, perhaps due to the energetics required to produce some high-magnitude signals, then signalers might evolve to stop producing the original component. When this occurs, the original component could be dropped completely from the species' repertoire, or it could be emancipated from the signal and co-opted for another function.

Assessing the phylogenetic history of signals among closely related species can reveal the loss of signals or signal components (Wiens 2001, Fig. 6.6). For example, several species of swordtail fish have dark bars on their sides, but some species have lost this visual signal component and the preference for it (Morris 1998). Wiens (2001) shows that trait loss commonly occurs when females switch their preference from one trait to another. A female preference swap can occur when one signal becomes more reliable than the other (such as in the context of making it possible to distinguish between conspecifics and heterospecifics), one signal becomes easier to detect because of changes in habitat or signal context, or one signal becomes increasingly costly (due to factors such as energetic requirements or predation risk) (Wiens 2001; Rosenthal et al. 2002). We suggest that the increasing difficulty of discrimination as a signal component increases in magnitude could be a common

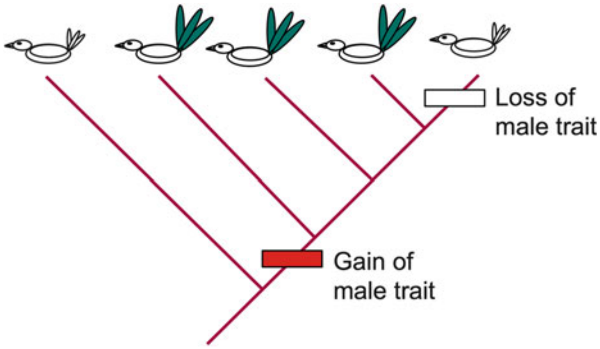


Fig. 6.6 Loss of original signal components. Phylogenies that show when a signal was gained by the common ancestor of several closely related species can reveal cases where it was then lost. (From Wiens 2001, permission granted by Elsevier)

reason that females switch to attend to another trait, resulting in loss of the original trait.

6.3.4 Directional Predictions

We describe above how directional sexual selection toward signal elaboration can be halted by proportional processing, but there are some communication systems for which we predict proportional processing to cause directional selection on signal evolution. For example, when two closely related species live in sympatry and signals from the two species have similar features, receivers can make errors in discrimination between conspecific and heterospecific calls (Howard 1993). When hybridization imposes a fitness cost on parents, one or both species might evolve signal changes through reproductive character displacement. These differences should make it easier for females to make the best mate choice decision and avoid hybridization. In this case, signal traits should move directionally away from each other, and proportional processing might influence the direction of movement (Akre and Johnsen 2014).

The factors that determine which species' signal moves and in which direction it moves are complex. Lemmon (2009) list relevant factors including differences between the two species in their postzygotic or prezygotic isolation, population densities, costs or constraints on trait evolution, and possible dilution of change due to a neighboring pool of allopatric individuals. When such asymmetries between the two species are slight or absent, however, female proportional processing could influence the direction of reproductive character displacement (Akre and Johnsen 2014). When proportional processing error in discrimination occurs, and all other factors remain the same, the movement of reproductive character displacement is more likely to occur in the direction of easier discrimination or toward lower-magnitude stimuli.

Suppose that two frog species live in sympatry: species A has a call with variation in pulse rate ranging from 6 to 10 pulses per second, and species B has call variation ranging from 8 to 12 pulses per second. The overlap will cause some receivers to choose heterospecifics as mates. If hybrids reduce these individuals' fitness, there will be selective pressure for reproductive character displacement or change in the signal and female preference. Will species A move lower, or will species B move higher? Considering an equal absolute change in both directions and assuming that change incurs the same costs in both directions, the benefits of moving lower would outweigh the benefits of moving higher, because receivers that discriminate by proportional processing would be more likely to notice the same absolute difference in the lower direction. This means the species with lower magnitude in the relevant signal parameter should be more likely to change, moving even lower in magnitude, when all other factors are equal. In reality there are many factors that can push reproductive character displacement in a particular direction, but considering discriminability could help explain some directional changes.

There are reported cases of sympatric species producing lower-magnitude call characteristics (Loftus-Hills and Littlejohn 1992), but other cases show movement toward higher magnitude, and some even show different directional solutions in distinct populations (Lemmon 2009).

The costs involved in signal production can dramatically alter this picture. In the scenario above, species B moving up an absolute increment is less beneficial than species A moving down by the same absolute increment. However, the cost of adding or subtracting absolute increments of signal magnitude can vary with magnitude, too. Both energetics and predation risk often make it costlier to increase than to decrease the magnitude of a signal. Making a vocal signal louder requires more effort. Making a pheromone stronger requires more synthesis. Making a visual signal bigger might require more movement. In all of these cases, a bigger signal could be more conspicuous to predators. These trends would further back up the idea that decreasing signal magnitude is more likely than increasing signal magnitude. However, the cost of producing a signal does not always increase with signal magnitude—for example, a species evolving a darker coloration might require the production of more melanocytes, while evolving a brighter coloration might occur through simply producing less melanocytes (Fig. 6.7). A study of reproductive character displacement in two damselfly species that differ in female wing brightness showed that in regions of sympatry, one of the species evolved brighter wing coloration relative to its allopatric populations (Waage 1975). Depending on the mechanism of wing coloration, the fact that a change toward greater brightness occurred could be partially affected by the costs of producing melanocytes. The predation cost of increasing a signal's magnitude could also decrease with signal magnitude in some situations. For example, the bioluminescent searchlights of deep sea predators are not effective when aimed at red prey that absorb the blue bioluminescent light. Cephalopods that dynamically control expansion of their red chromatophores reduce predation risk by increasing the magnitude of their chromatophore coverage (Zylinksi and Johnsen 2011).

Several behavioral contexts other than evolving reproductive character displacement might similarly influence directional signal evolution to improve receiver discriminability. For example, some animals change the frequencies of their vocalizations in response to noise in the environment (Hu and Cardoso 2010). Animals that face the introduction of a new source of narrowband noise in the environment might improve their communication efficacy by evolving directionally away from the range of magnitudes characteristic of the noise. And animals that face parasitic mimicry might evolve away from their original signal value to escape the costs of parasitism. In each case, the factors determining directional movement will be complex, but if costs are equal in both directions, proportional processing could tip the scales to move toward the lower magnitude, because the benefits of receiver discrimination will be greater for the same absolute change.

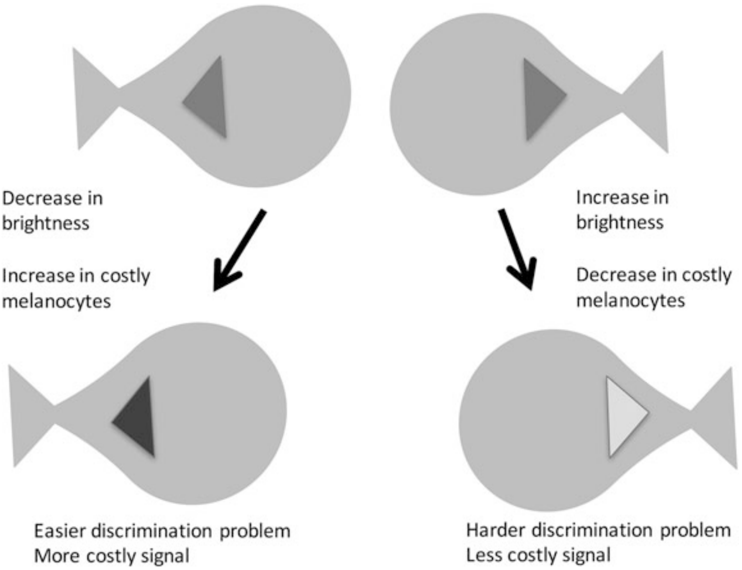


Fig. 6.7 Expected relationships between signal magnitude, signal costs, and signal discrimination can be decoupled in some cases. If two closely related fish species both use a brightness patch to recognize potential mates but have a very similar patch brightness, reproductive character displacement might occur such that one species evolves a change in brightness in regions of sympatry. Based on proportional processing, receivers would more easily note the same absolute amount of change if the signal evolves toward lower magnitude or decreases in brightness. However, the cost of decreasing signal magnitude is greater than the cost of increasing magnitude in this case. Producing more melanocytes would be costlier than reducing melanocyte production. Thus, proportional processing is unlikely to influence the direction of signal change in this case

6.4 Impacts on Signaler Strategies

Since signalers communicate to influence the behavior of receivers, receiver errors resulting from proportional processing can influence the evolution of strategies for signal display. Strategic production of communication signals is possible when signals are used in dynamic displays (e.g., Chap. 10). For example, males advertising for females might display strategically according to the physical and competitive landscape (Lindstrom et al. 2009, Wong and Candolin 2005; Goutte et al. 2010). And strategic variation in display might occur in deceptive situations, such as the behavioral mimicry shown by some brood parasite hatchlings (Langmore et al. 2003). In this section, we discuss how proportional processing by receivers could influence signaler strategy on a dynamic time scale. Many of these effects parallel the influence on signal properties that occurs on an evolutionary time scale.

6.4.1 Parallels to Evolutionary Consequences of Proportional Processing

In our discussion of how proportional processing influences the evolution of signal design, we showed how receiver proportional processing might limit the evolution of increasingly elaborate sexual signals. This occurs when the benefit of increasing magnitude diminishes with each increment, because female receivers cannot perceive the additional increments. This same process should influence signaling strategy on a much shorter time scale—within a bout of mate advertisement. Due to female inability to notice additional increments at higher magnitude, the benefit of increasing magnitude will diminish as signals increase in magnitude. Thus, males that can dynamically control the magnitude of their signals should strategically reduce their probability of increasing signal magnitude once they reach high magnitudes, even if they are physically capable of increase. Generally, costs increase with increasing magnitude, due to energy requirements and increased conspicuousness. The combined effect of increasing costs and reduced benefits should place strong selective pressure on males to evolve strategies that limit dynamic use of highly elaborate signals. Túngara frog males exhibit this type of strategy—the probability they will add call elements decreases as the overall number of elements increases (Fig. 6.8; Akre and Ryan 2011).

Competing males could consider an alternative solution to females' inability to discriminate between high-magnitude signals that are close to each other in magnitude. If a male were to suddenly increase the magnitude of his signal dramatically, for example, upon detecting a female, the difference between signals might be easier for females to detect. This appears to be a signaling strategy adopted by some male frogs. As illustrated in Fig. 6.8, a male túngara frog is more likely to produce simple calls lacking chucks (i.e., "whines" only) in the absence of female movement (cf. Fig. 6.8a, b; 0.82 versus 0.56). However, when a nearby female's movement is detected, males become more likely to add one or two chucks to his calls (cf. Fig. 6.8a, b; 0.44 versus 0.18 and 0.35 versus 0.13). Males of the gray treefrog (*H. versicolor*) increase the duration of their vocal signal by adding pulses in response to detecting the presence of a female using either tactile (Reichert and Gerhardt 2012) or visual (Reichert 2013) sensory perception. The resulting dramatic increase in pulse number upon detecting a female might make their signal stand out, in spite of proportional processing, by helping females detect a difference even at high magnitude.

Another parallel to evolutionary consequences of proportional processing is facultatively adding distinct components to a display only when the benefits are most likely to occur. If adding a second type of component is costly due to energetics or predation, an advertising male could withhold the signal until the first component is competitively ineffective. As the comparison of male signals becomes harder with competitive escalation, males should strategically add new axes of signal variation into their display, giving females another chance to compare potential mates. Adding a secondary display component would be worth

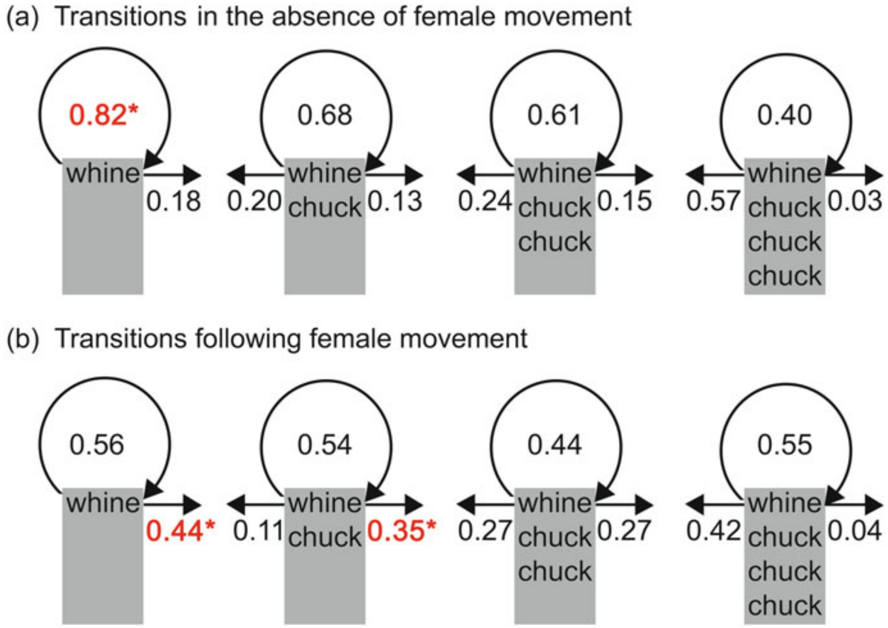


Fig. 6.8 Strategic modifications of signal magnitude. Male túngara frogs (*Physalaemus pustulosus*) strategically decide when to transition between calls varying in chuck number in ways expected for proportional processing by female receivers. At any moment, males can decrease, maintain, or increase the number of chucks appended to their simple call (the "whine"). Shown here are transition probabilities between signals differing in signal magnitude (i.e., chuck number) in (**a**) the absence and (**b**) the presence of female movement. In the absence of female movements, males are relatively more likely to produce consecutive whines than when following female movement (0.82 versus 0.56). In contrast, males become more likely to transition to calls having one or two chucks following female movement (0.44 versus 0.18 and 0.35 versus 0.13). Note, additionally, that in both cases, males become less likely to transition to a call with more chucks as the absolute magnitude of their signal (chuck number) increases. These data are consistent with expectations based on proportional processing of signal magnitude (cf. Fig. 6.3). *Bold red font with asterisks* indicates significant differences between the two female movement conditions. (Modified from Akre and Ryan 2011, permission granted by Oxford University Press)

the added costs once one axis gets maxed out and a male can no longer outcompete other males. This scenario parallels the evolutionary changes in signal properties that could result from receiver proportional processing but occurs within a single individual's flexible behavioral repertoire.

In competitions, opponents must track each other's escalation, and again increasing magnitudes would be hard to detect at high intensity. One way to solve this problem is to reverse the direction of escalation, toward decreasing magnitude. For some animals, contest and attraction are communicated in one signal, but animals can evolve a way to separate these signals. In gray treefrogs (*H. versicolor*), males in aggressive confrontations use calls with lower dominant frequency than the calls used to attract females. In this case, as competition escalates, call frequency decreases (Reichert and Gerhardt 2013). This change is costly to males but would make it easier to determine the winner of a contest using frequency as the determining factor.

6.4.2 Trade-offs Between Pushing High-Energy Signals and Maintaining Steady Mid-level Signals

When individuals assess signal magnitude over a long period of signaling in order to choose a mate or make a decision about some other resource, proportional processing can make them remember the most likely signal magnitude as lower than it actually is. This is because cognitive calculations of which signal is most likely to occur depend on perception of signal magnitude, which can only be as accurate as the JND. Thus, as a single signal magnitude occurs repeatedly, it can actually be perceived at a distribution of values within the JND, and the probability that a signal will occur at a given point is diluted. Since JNDs are larger at higher magnitudes, the probability of any one quantity gets especially diluted at high magnitude, as the possible perceived quantities occur within a wider window (Bateson and Kacelnik 1995). Thus, when perception occurs in this way, the likelihood of any one magnitude will be exaggerated for lower magnitudes relative to higher one (Fig. 6.9). It is possible that for some sensory parameters, repeated signals are not actually assessed individually but are instead considered as an identical group, ignoring slight perceptual differences. But a false memory of lower magnitude does occur sometimes. This effect has been shown to influence the choices made by European starlings (*Sturnus vulgaris*) about variable versus fixed amounts of food, such that they choose fixed rewards over variable ones even when the physical average amount is equal for both options and variable delays over fixed delays when they average to the same delay (Reboreda and Kacelnik 1991).

In any situation in which animals use long strings of communication signals to assess value, a receiver's decision could be influenced by this proportional processing-based bias. For example, if females choose a mate by assessing a string of signals produced by a male and determining the most likely signal produced, her assessment will be biased toward lower values. Some receivers will assess a different aspect of repeated signaling, such as rejecting if the lowest signal is below a threshold or accepting if the highest signal is above a threshold. But when most likely signal value is assessed, proportional processing could influence the evolution of male display strategies. A female choosing a male based on the magnitude of his most likely signal will prefer a male that produces a string of average magnitude signals over a male who produces an equal amount of high-magnitude signals and low-magnitude signals. In fish that produce tail fans that attract females, the tail fan can indicate direct benefits from parental care to come (Pampoulie et al. 2004). If males fan their tails to oxygenate eggs, females choosing

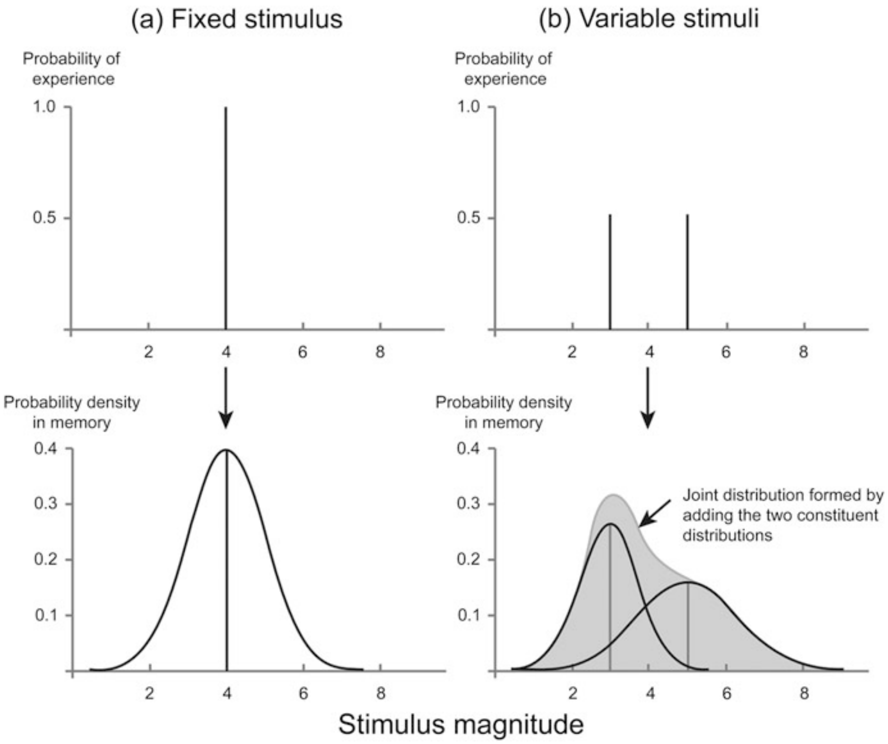


Fig. 6.9 The influences of proportional processing on an animal's perception of fixed versus variable stimuli repeated over time. (**a**) When animals experience repeated exposure to a fixed stimulus, depicted here as a stimulus of magnitude 4 (**a**, *top*), perceived stimulus magnitudes are distributed within the window of the JND. This creates variability in the animal's memory of stimulus magnitude, depicted as a probability density function (**a**, *bottom*). As illustrated here, for fixed stimuli, the peak probability in the distribution of perceived magnitudes corresponds well to the actual signal magnitude. Thus, repeated exposure to a fixed or invariant stimulus can result in good correspondence between actual and perceived stimulus magnitudes in memory. (**b**) When repeated stimuli are more variable, depicted here as stimuli produced with equal probability at two different magnitudes (3 and 5; **b**, *top*), animals may remember the lower-magnitude variant as being more common than it actually was. This is because the perceived magnitude of the low-magnitude and high-magnitude stimulus variants are each distributed within the window of the JND, which, as a proportion of stimulus magnitude, corresponds to a larger absolute difference for the higher-magnitude stimulus (**b**, *bottom*). Hence, the distribution of perceived magnitudes for higher-magnitude stimuli is necessarily more variable as a result of proportional processing. According to the joint distribution of perceived stimulus magnitudes (*shaded area* in **b**, *bottom*), animals would perceive the lower-magnitude signal as being more common than the higher-magnitude signal. (From Bateson and Kacelnik 1995, permission granted by John Wiley and Sons, Inc.)

males based on most likely tail fanning rate are likely choosing males that will fan eggs effectively. A male strategizing to attract a female based on tail fanning rate might decide to put all his energy into a burst of fast fanning or to spread out his energy and produce a long stream of medium fanning. Based on females remembering likely fanning rates with a bias toward lower values, the male should produce a long stream of medium fanning rates rather than a quick burst of fast fans followed by a long stream of slow fans. This type of assessment could also impact males determining how to defend their territories when signaling over a long period of time.

This has not been tested in many natural settings. However, gray treefrog (*Hyla versicolor*) females do not discriminate between highly variable and low variability strings of advertisement calls varying in duration, call period, or both when the means are the same or nearly the same (Gerhardt and Watson 1995). They do, however, discriminate when the mean is slightly different, indicating that they can keep track of differences between individuals over extended periods of calling (Gerhardt and Watson 1995; Schwartz et al. 2004).

6.5 Impacts on Receiver Responses

Proportional processing influences the evolution of signals and signalers in predictable ways because it asymmetrically compromises a receiver's ability to accurately detect and categorize signals, with greater error for higher-magnitude signals. Sometimes, however, receivers will evolve a way to reduce the negative impacts of proportional processing-based errors. All types of sensory perception—not just proportional processing—involve JNDs that keep an individual from precisely translating physical stimulus value into a perceived value. The impact of any type of JND-related error can be reduced if animals evolve smaller JNDs. But mechanistically decreasing the JND could be costly and reach a point where physical limitations prevent further reduction. For example, visual contrast discrimination can be improved by evolving a larger eye that lets in more light or by suffering a reduction in acuity (Cronin et al. 2014), but both are costly changes.

When receivers that use proportional processing face the problem of JND-based errors, they could evolve an alternative strategy to use specifically when assessing the high-magnitude signals that result in more perceptual errors. Such a strategy differs from other types of perception that encounter problematic JNDs equally at any stimulus magnitude, or when JNDs change across the range of detectable stimulus magnitudes, but in a manner inconsistent with proportional processing. Whenever distinct strategies are appropriate for different ranges of stimulus magnitude, each should be maintained in a receiver's behavioral repertoire. Thus, proportional processing could generate behavioral diversity and behavioral flexibility within receivers. In this section, we describe two behaviors that receivers could use in the context of high-magnitude signals to reduce the impact of proportional processing-based error: shifting attention toward different sensory parameters or signal components and producing a compensatory behavior to improve response accuracy. We then discuss how a consideration of proportional processing impacts signal detection theory models of receiver decision-making.

6.5.1 Attention to New Sensory Parameters or Signal Components

The JNDs for some sensory inputs are more likely than others to cause categorization error. Since receiver processing should evolve to maximize fitness, receivers could evolve to ignore those parameters that are hard to measure accurately. Which parameters are more accurate depends on several factors. Some parameters have a small JND relative to behaviorally relevant signal variation, making behaviorally relevant differences between signals easier to detect. Radiance signals, for example, are generally produced at magnitudes far above the contrast threshold, making them easy to detect (Cronin et al. 2014). But some sensory parameters have a large JND relative to behaviorally relevant variation, so behaviorally relevant differences are difficult to detect. Those sensory parameters that fall in this last category are the ones that receivers might evolve a way to ignore.

Since receivers have access to several sensory parameters when assessing a signal, they could evolve a change in assessment strategy, shifting from one parameter to another. For example, a vocalization can be processed in terms of frequency, amplitude, duration, or rate, among other possibilities. If signal duration evolves increasing magnitude to the point that a receiver cannot distinguish between two signal durations, receivers could evolve to attend instead to differences in rate of production. Of course, there would be no evolutionary selection for shifting if doing so had no benefit to the receiver. Thus, this is most likely to occur when signal assessment truly impacts receiver fitness, for example, when signals are honest indicators of some signaler characteristic.

The feasibility of this solution depends on whether the new candidate parameter has the potential to function the same way the original parameter did. For example, females assessing a signal component that indicates good genes might prefer elaborate versions of this component, causing it to evolve to the point where females cannot distinguish between males. Determining when a lack of preference is the result of inability to discriminate is of course complicated by the possibility that excessive exaggeration is behaviorally irrelevant; this reflects the JND versus JMD issue we discussed in Sect. 6.2. But when sensory constraints are influencing behavior, attending to a different component of the signal might be a good strategy if it, too, reliably predicts good genes. In gray treefrogs, call duration indicates good genes (Welch et al. 1998). If females assess duration temporally and become unable to distinguish between long call durations, changing their attention toward fundamental frequency or peak call amplitude might not correlate with good genes in the same way. However, if they assess duration temporally, changing their attention toward total call energy or a count of pulses would rely on a different sensory parameter with a potentially different Weber fraction (k). This assessment could also become a difficult discrimination problem at high magnitude, but this k might not constrain females at the relevant range of magnitudes. Shifting attention from duration toward call rate (or call period) might also work if call rate correlates with call duration (Wells and Taigen 1986) and call rate (or call period) is assessed along a different sensory parameter with a different k. If the k for a distinct sensory parameter allows for more accurate comparisons at the appropriate range of magnitude, this option could allow females to discriminate the calls representing better genes even with heightened elaboration.

When shifting sensory parameters is a valid option, receivers that use proportional processing could strategically assess high-magnitude communication signals differently than low-magnitude signals. For example, a honeybee perceiving a returned forager dancing the waggle dance (Fig. 6.1) could switch to a new strategy for assessing waggle duration when the food source being danced about is especially far away, introducing the possibility of larger absolute error. She could switch between assessing waggle duration using visual, tactile, or acoustic information if the discrimination threshold for one parameter was finer than that for the others. Or, she could verify an uncertain signal assessment by calling upon more than one sensory modality (see also Chaps. 4, 5 and 11). The solution determined by a receiver will depend on the variety of available sensory channels for information flow and the costs associated with each. Few studies directly assess this question. However, there is evidence that ants with larger discrimination thresholds recruit more sensory modalities in navigation decision-making. Von Thienen et al. (2014) found that of three species of ants using pheromone concentration to decide on which pathway to take, the species with the largest discrimination threshold for olfactory concentration also uses visual information to navigate. This could be the result of evolution to reduce receiver error.

6.5.2 Evolving a Compensatory Behavior

When it is costly or ineffective to evolve a better discrimination threshold or switch attention between sensory parameters, receivers might evolve compensatory behaviors to make up for error in their perception of communication signals. Compensatory behaviors are likely to occur whenever receiver error is costly, not just when error is caused by proportional processing. For example, mothers in the superb fairy wren (*Malurus cyaneus*) faced with the inability to distinguish between her own and mimetic brood parasitic eggs have evolved incubation calls that only her own hatchlings have the opportunity to learn. This way, she can test hatchlings' knowledge of the password before investing in the effort of raising them (Colombelli-Négrel et al. 2012). When the difficulty distinguishing between or recognizing signals is the result of proportional processing, however, compensatory behaviors are likely to be used specifically in response to higher-magnitude stimuli rather than all the time.

Communication about navigation is an example of a context in which receivers are likely to evolve compensatory behavior for use in high-magnitude stimulus assessment specifically. When animals translate the perceived value of communication signals into traveling a physical distance, as occurs with the honeybee's waggle dance (Fig. 6.1), it is likely that their error will increase with the distance traveled. Using multiple strategies when the potential error is high would help them find their end point goal. Honeybees can use olfactory information in this way. When olfactory information is withheld from navigating honeybee foragers that have watched the dance of another forager, they exhibit searching behavior and are usually unable to find the exact location of their target (Riley et al. 2005; Tautz and Sandeman 2003). Thus, the use of olfactory cue-based searching could be a compensatory behavior in honeybees.

6.5.3 Proportional Processing and Signal Detection Theory

Considering how proportional processing impacts receiver response to signals could improve our understanding of *signal detection theory* (*SDT*). SDT, as applied to animal communication, describes how animals can best respond to potential error in signal categorization (Green and Swets 1966; Wiley 2006, 2013). Animals evolve to respond adaptively to stimuli in the environment. But knowing which stimuli correspond to a particular entity in the environment, such as a signal produced by a conspecific rather than by a heterospecific, is not always easy. Two stimuli from very different sources can be very similar (Caldwell et al. 2009). Environmental factors can degrade signal quality, making the distinction harder (Cronin et al. 2014; Wiley and Richards 1978). Two main features of SDT describe how animals address this problem: setting a response threshold (a binary situation, responding or not responding, or for classification, responding as A or B) and realizing four possible outcomes to any response—correct yes, incorrect yes, correct no, and incorrect no (Wiley 2013; Chap. 7). Each of these four outcomes has its own costs and benefits associated with it. The relative costs and benefits of each should determine the threshold value of perceived signal variation that determines the point at which animals will shift their binary response. If the costs for an incorrect yes response are very high, animals will only respond when the stimulus is well within the range of variation that indicates high probability for the target stimulus. If the costs of an incorrect yes response are very low, the range that triggers a response will be much broader and overlap more with the nontarget stimulus range.

Errors in receiver perception complicate the determination of where the most beneficial threshold value lies. Some receiver error is not consistent enough to model, such as distraction by other interactions or internal state changes. But the errors caused with consistent patterns, such as by JNDs, can be added to the model of receiver response options. One impact of proportional processing in particular would be a redistribution of the probability density functions that are used in the calculations of SDT. SDT models the probability that any given value comes from the signal of interest and the probability that it comes from an alternative source—noise or a similar signal. Proportional processing should distort the shape of these SDT curves. With proportional processing, receiver error is more likely as magnitude increases. Thus, the curve of probability for a given signal should slightly spread out in the direction of higher magnitude. This change might slightly shift the threshold value determined to be the optimal point of shifting response for an animal trying to maximize correct response and minimize costs. The influence of proportional processing on SDT is minimal but possibly helpful in understanding when behavior does not match standard SDT predictions precisely.

6.6 When JNDs Follow Other Patterns of Change with Variation in Stimulus Magnitude

The reason we use proportional processing in particular as a point from which to make predictions about the evolution of behavior is because this type of sensory processing occurs over at least part of the range of detectable magnitudes in many sensory systems in many taxa. It applies to the visual parameters of area and brightness, the acoustic parameters of loudness and frequency, the olfactory or gustatory assessment of concentration, the tactile assessment of pressure, assessments of number or time, and even some electrosensory parameters (Akre and Johnsen 2014). It has been documented in invertebrates, fish, anurans, reptiles, birds, and mammals (Akre and Johnsen 2014). Thus, proportional processing-based studies can be applied broadly. But there are many departures from the patterns predicted by Weber's law, and some sensory parameters, such as color vision and magnetoreception, do not follow Weber's law at all.

In cases where the transformation between actual and perceived stimulus values do not have a consistent proportional JND, more complicated patterns of signal categorization can be predicted based on similar principles. For these sensory parameters, JNDs still sometimes shift in regular ways across the range of detectable magnitudes. An example of this is wavelength discrimination in color vision. Humans can discriminate certain regions of the spectrum far better than others (Cronin et al. 2014). Another example is brightness discrimination. Human brightness perception has a consistent Weber fraction (k) at diurnal light levels, but at lower light levels it follows another pattern (the DeVries-Rose law), and at very high light levels a third pattern occurs—the proportional JND increases with intensity (Cronin et al. 2014). Patterns such as this could lead to new sets of concrete predictions about the evolution of communication. If a proportional JND increases with intensity rather than stays at a constant proportion, the effects of proportional processing outlined in this chapter would be even more dramatic. That is, the absolute increase in a JND, when moving to higher signal magnitudes, would be an even greater absolute increase.

We encourage researchers to consider these alternative patterns of change in studies that depend upon response to these ranges of magnitude or these other sensory parameters. For example, studies of anurans and insects demonstrate an "opposite miss" to Weber's law (Gerhardt 1987; Forrest 1994; Wyttenbach and Farris 2004; Bee et al. 2012), as discussed in Sect. 6.2. The result is even worse discrimination ability than predicted by proportional processing as sounds reach higher amplitude. Amplitude can influence female phonotaxis in several ways (Gerhardt 1981, Beckers and Schul 2004). The “opposite miss” factor indicates that it should have less impact at close proximity. In fact proximity can change phonotaxis behaviors in anurans (Akre and Ryan 2010) and insects (Mendelson and Shaw 2006), and this psychophysical perspective indicates that females will place greater weight on non-amplitude features of calling in close-range decisions. Birds, however, show a “near miss” to Weber’s law, with increased sensitivity to amplitude at close range. Therefore, they might continue to use amplitude as a relevant vocalization feature even when callers escalate to very high amplitudes. Together, these studies illustrate the potential for interesting evolutionary differences across taxa that could differently impact the evolution of their communication systems.

6.7 Summary and Future Directions

This chapter has demonstrated that understanding how sensory systems translate stimulus magnitude into perceived magnitude can illuminate selective forces in the evolution of animal communication behaviors. Proportional processing is one such pattern of translation that results in a predictable asymmetry in the perception of higher versus lower-magnitude stimuli. This particular pattern is widespread across certain ranges of many sensory parameters in many animal taxa. When receivers use this type of processing, we can predict how signals might evolve exaggeration, new components, loss of components, and directional change. We can predict when signalers might strategize to accelerate or change their displays or strategize a string of signals most likely to result in a favorable receiver response. We can predict when receiver sensory systems will evolve versus when receiver behaviors will evolve in terms of attention allocation or adding novel behaviors to compensate for receiver errors. These predictions only apply when animals follow Weber’s law, but when other patterns occur in different sensory parameters or different ranges of magnitude, we could devise more specific predictions for those cases based on the principles used in this chapter. Of course, the translation from actual to perceived signals is just one of many factors that impact the evolution of animal communication, but it is an important part of understanding receiver psychology that has not been investigated previously in much depth.

Future work should test more of these predictions to determine when proportional processing influences signal categorization error and the evolution of communication. This would begin by identifying where proportional processing can influence behavior and then predicting how this might influence behavior. Producing models of what behavior influenced by proportional processing would look like will allow comparison with the actual behavior response observed. However, determining whether a match between observed and predicted behavior based on models of proportional processing truly results from JNDs will require ruling out other possibilities, as discussed in Sect. 6.2.

Previous work has identified a few ways to begin approaching this challenge (Akre et al. 2011; Akre and Johnsen 2014). Models of the behavioral responses that would occur if animals are constrained by JNDs can be compared to other models based on the possible behavioral significance of the signal variation. For example, potential predation threat could increase with predator proximity, so response to predator vocalizations might follow a curve based on increasing threat rather than insensitivity to amplitude change. Finding a way to quantify how predation threat varies with proximity would allow testing whether this model fits the actual response data better than the model based on proportional processing. Or another example is that larger males might produce more sperm. By finding a way to quantify the relationship between body size and sperm production, one could test whether this model explains actual female response patterns better than a model based on her proportional processing of male body size. Eliminating these obvious alternative possibilities increases the probability that proportional processing is responsible for behavioral patterns.

Addressing the challenge of differentiating responses based on JND from those based on JMD could benefit from a focus on Weber fractions (k), the proportional constant that is used in the equation for Weber's law. If one species responds to stimuli with completely different meanings in a manner reflecting the same underlying k value, it is more likely that responses result from proportional processing limitations. For example, if a species responds to two separate signals within the same sensory modality (and parameter), a response guided by JNDs will have the same JND regardless of the function of that signal. Thus, a fish responding to visual assessment of increasing size of predatory threat and potential prey will show the same k guiding response curves if the response is constrained by JNDs. Perhaps more difficult to evaluate is a receiver's lack of response to a specific difference in signal magnitude, especially in studies that measure natural responses to natural stimuli. Did the receiver fail to respond because the magnitude difference failed to exceed its JND? Or was the JND exceeded (meaning the receiver perceived the difference) but the difference was smaller than its JMD (meaning the difference was too small to warrant a discriminative behavioral response)? Finding answers to such questions may require the use of classical or operant conditioning techniques, which may allow experimenters to exert more overt stimulus control over receiver behavior (see, e.g., Chap. 3).

We encourage future researchers to consider broad applications of this type of thinking. For example, applying a study of proportional processing to the evolution and development of human language could generate new and valuable ideas. The incredible explosion of innovation generating variability in language might stem in part from the futility and inefficiency of discriminating minor changes in magnitude along one axis of variation. Also, looking for relationships between how proportional processing impacts animal communication and how it impacts related cognitive faculties such as music, math, or emotion, might be valuable (Akre and Johnsen 2014). We hope that considering the salience of proportions in animal perception and cognition will generate productive new directions of thought and research on receiver psychology and the evolution of animal communication.

Acknowledgments We thank Mark Bee, Cory Miller, and an anonymous reviewer for their helpful comments on the manuscript. We thank Melissa Bateson, Alex Kacelnik, and John Wiens for the use of their excellent figures.

References

Akre KL, Johnsen S (2014) Psychophysics and the evolution of behavior. Trends Ecol Evol 29:291–300

Akre KL, Ryan MJ (2010) Proximity-dependent response to variably complex mating signals in túngara frogs (*Physalaemus pustulosus*). Ethology 116(12):1138–1145

Akre KL, Ryan MJ (2011) Female túngara frogs elicit more complex mating signals from males. Behav Ecol 22:846–853

Akre KL, Farris HE, Lea AM, Page RA, Ryan MJ (2011) Signal perception in frogs and bats and the evolution of mating signals. Science 333:751–752

Andersson M (1994) Sexual selection. Princeton University Press, Princeton, NJ

Andersson M, Simmons LW (2006) Sexual selection and mate choice. Trends Ecol Evol 21:296–302

Bateson M, Kacelnik A (1995) Preferences for fixed and variable food sources: variability in amount and delay. J Exp Anal Behav 63:313–329

Beckers OM, Schul J (2004) Phonotaxis in *Hyla versicolor* (Anura, Hylidae): the effect of absolute call amplitude. J Comp Phys A 190(11):869–876

Bee MA (2008) Parallel female preferences for call duration in a diploid ancestor of an allotetraploid treefrog. Anim Behav 76:845–853

Bee MA, Gerhardt HC (2002) Individual voice recognition in a territorial frog (*Rana catesbeiana*). Proc R Soc Lond B 269(1499):1443–1448

Bee MA, Vélez A, Forester JD (2012) Sound level discrimination by gray treefrogs in the presence and absence of chorus-shaped noise. J Acoust Soc Am 131(5):4188–4195

Beecher MD, Petersen MR, Zoloth SR, Moody DB, Stebbins WC (1979) Perception of conspecific vocalizations by *Japanese macaques*. Brain Behav Evol 16(5–6):443–460

Bernal XE, Page RA, Rand AS, Ryan MJ (2007) Natural history miscellany: cues for eavesdroppers: do frog calls indicate prey density and quality? Am Nat 169:409–415

Bernal XE, Akre KL, Baugh AT, Rand AS, Ryan MJ (2009) Female and male behavioral response to advertisement calls of graded complexity in túngara frogs, *Physalaemus pustulosus*. Behav Ecol Sociobiol 63:1269–1279

Bullock TH, Hopkins CD, Fay RR (eds) (2006) Electroreception, vol 21. Springer Science & Business Media, New York, NY

Bush SL, Gerhardt HC, Schul J (2002) Pattern recognition and call preferences in treefrogs (Anura: Hylidae): a quantitative analysis using a no-choice paradigm. Anim Behav 63:7–14

Caldwell MS, McDaniel JG, Warkentin KM (2009) Frequency information in the vibration-cued escape hatching of red-eyed treefrogs. J Exp Biol 212(4):566–575

Charrier I, Mathevon N, Jouventin P (2002) How does a fur seal mother recognize the voice of her pup? An experimental study of *Arctocephalus tropicalis*. J Exp Biol 205(5):603–612

Cohen JA (1984) Sexual selection and the psychophysics of female choice. Z Tierpsychol 64:1–8

Colombelli-Négrel D, Hauber ME, Robertson J, Sulloway FJ, Hoi H, Griggio M, Kleindorfer S (2012) Embryonic learning of vocal passwords in superb fairy-wrens reveals intruder cuckoo nestlings. Curr Biol 22:2155–2160

Cronin TW, Johnsen S, Marshall NJ, Warrant EJ (2014) Visual ecology. Princeton UniversityPress, Princeton, NJ

Endler JA, Westcott DA, Madden JR, Robson T (2005) Animal visual systems and the evolution of color patterns: sensory processing illuminates signal evolution. Evolution 59:1795–1818

Fechner GT (1860) Elemente der Psychophysik (2 vols). English edition: Fechner GT (1966) Elements of psychophysics, vol 1 (trans: Adler HE). Breitkopf and Hartel

Forrest TG (1994) From sender to receiver: propagation and environmnetal effects on acoustic signals. Am Zool 34:644–654

Gabor CR, Ryan MJ (2001) Geographical variation in reproductive character displacement in mate choice by male sailfin mollies. Proc R Soc Lond B 268:1063–1070

Gerhardt HC (1981) Mating call recognition in the green treefrog (*Hyla cinerea*): importance of two frequency bands as a function of sound pressure level. J Comp Physiol 144(1):9–16

Gerhardt HC (1987) Evolutionary and neurobiological implications of selective phonotaxis in the green treefrog, *Hyla cinerea*. Anim Behav 35(5):1479–1489

Gerhardt HC (1994) Reproductive character displacement of female mate choice in the grey treefrog, *Hyla chrysoscelis*. Anim Behav 47:959–969

Gerhardt HC, Watson GF (1995) Within-male variability in call properties and female preference in the grey treefrog. Anim Behav 50:1187–1191

Gerhardt HC, Tanner SD, Corrigan CM, Walton HC (2000) Female preference functions based on call duration in the gray tree frog (*Hyla versicolor*). Behav Ecol 11(6):663–669

Goutte S, Kime NM, Argo TF, Ryan MJ (2010) Calling strategies of male túngara frogs in response to dynamic playback. Behaviour 147:65–83

Green DM, Swets JA (1966) Signal detection theory and psychophysics. WiIey, New York, NY, London, Sydney

Guilford T, Dawkins MS (1991) Receiver psychology and the evolution of animal signals. Anim Behav 42:1–14

Howard DJ (1993) Reinforcement: origin, dynamics, and fate of an evolutionary hypothesis. In: Harrison RG (ed) Hybrid zones and the evolutionary process. Oxford University Press, Oxford, pp. 46–69

Hu Y, Cardoso GC (2010) Which birds adjust the frequency of vocalizations in urban noise? Anim Behav 79(4):863–867

Jennions MD, Petrie M (1997) Variation in mate choice and mating preferences: a review of causes and consequences. Biol Rev 72:283–327

Langmore NE, Hunt S, Kilner RM (2003) Escalation of a coevolutionary arms race through host rejection of brood parasitic young. Nature 422(6928):157–160

Lemmon EM (2009) Diversification of conspecific signals in sympatry: geographic overlap drives multidimensional reproductive character displacement in frogs. Evolution 63(5):1155–1170

Lindstrom J, Pike TW, Blount JD, Metcalfe NB (2009) Optimization of resource allocation can explain the temporal dynamics and honesty of sexual signals. Am Nat 174:515–525

Loftus-Hills JJ, Littlejohn MJ (1992) Reinforcement and reproductive character displacement in *Gastrophryne carolinensis* and *G. olivacea* (Anura: Microhylidae): a reexamination. Evolution 46:896–906

Mendelson TC, Shaw KL (2006) Close-range acoustic signaling and mate choice in Hawaiian crickets (Gryllidae: Laupala). Behav Ecol Soc 59(6):770–776

Miller CT, Bee MA (2012) Receiver psychology turns 20: is it time for a broader approach? Anim Behav 83:331–343

Morris MR (1998) Further examination of female preference for vertical bars in swordtails: preference for 'no bars' in a species without bars. J Fish Biol 53:56–63

Nelson DA, Marler P (1990) The perception of birdsong and an ecological concept of signal space. In: Stebbins WC, Berkley MA (eds) Comparative perception, vol 2: complex signals, Wiley series in neuroscience. Wiley, Oxford

Page RA, Ryan MJ (2008) The effect of signal complexity on localization performance in bats that localize frog calls. Anim Behav 76:761–769

Pampoulie C, Lindström K, St Mary CM (2004) Have your cake and eat it too: male sand gobies show more parental care in the presence of female partners. Behav Ecol 15:199–204

Reboreda JC, Kacelnik A (1991) Risk sensitivity in starlings: variability in food amount and food delay. Behav Ecol 2:301–308

Reichert MS (2013) Visual cues elicit courtship signals in a nocturnal anuran. Behav Ecol Sociobiol 67(2):265–271

Reichert MS, Gerhardt HC (2012) Trade-offs and upper limits to signal performance during close-range vocal competition in gray tree frogs *Hyla versicolor*. Am Nat 180(4):425–437

Reichert MS, Gerhardt HC (2013) Gray tree frogs, *Hyla versicolor*, give lower frequency aggressive calls in more escalated contests. Behav Ecol Sociobiol 67:795–804

Riley JR, Greggers U, Smith AD, Reynolds DR, Menzel R (2005) The flight paths of honeybees recruited by the waggle dance. Nature 435:205–207

Rosenthal GG, Wagner WE, Ryan MJ (2002) Secondary reduction of preference for the sword ornament in the pygmy swordtail *Xiphophorus nigrensis* (Pisces: Poeciliidae). Anim Behav 63:37–45

Rowe C (1999) Receiver psychology and the evolution of multicomponent signals. Anim Behav 58:921–931

Rowe C (2013) Receiver psychology: a receiver's perspective. Anim Behav 85:517–523

Ryan MJ (1985) The túngara frog, a study in sexual selection and communication. University of Chicago Press, Chicago, IL

Ryan MJ (1998) Sexual selection, receiver biases, and the evolution of sex differences. Science 281:1999–2003

Ryan MJ, Rand AS (1993) Species recognition and sexual selection as a unitary problem in animal communication. Evolution:647–657

Schwartz JJ, Buchanan BW, Gerhardt HC (2001) Female mate choice in the gray treefrog (*Hyla versicolor*) in three experimental environments. Behav Ecol Sociobiol 49(6):443–455

Schwartz JJ, Huth K, Hutchin T (2004) How long do females really listen? Assessment time for female mate choice in the grey treefrog, *Hyla versicolor*. Anim Behav 68(3):533–540

Servedio MR, Noor MAF (2003) The role of reinforcement in speciation: theory and data. Annu Rev Ecol Evol Syst 34:339–364

Tautz J, Sandeman DC (2003) Recruitment of honeybees to non-scented food sources. J Comp Phys A 189:293–300

von Frisch K (1967) The dance language and orientation of bees. Harvard University Press, Cambridge, MA

von Thienen W, Metzler D, Choe DH, Witte V (2014) Pheromone communication in ants: a detailed analysis of concentration-dependent decisions in three species. Behav Ecol Sociobiol 68:1611–1627

Waage JK (1975) Reproductive isolation and the potential for character displacement in the damselflies *Calopteryx maculata* and *C. aequabilis* (Odonata: Calopterygidae). Syst Zool 24:24–36

Weber EH (1834) De Pulsu, resorptione, auditu et tactu: annotationes anatomicae et physiologicae. Koehlor, Lerpzig

Welch AM, Semlitsch RD, Gerhardt HC (1998) Call duration as an indicator of genetic quality in male gray tree frogs. Science 280:1928–1930

Wells KD, Taigen TL (1986) The effect of social interactions on calling energetics in the gray treefrog (*Hyla versicolor*). Behav Ecol Sociobiol 19:9–18

Wiens JJ (2001) Widespread loss of sexually selected traits: how the peacock lost its spots. Trends Ecol Evol 16:517–523

Wiley RH (2006) Signal detection and animal communication. Adv Stud Behav 36:217–247

Wiley RH (2013) Signal detection, noise, and the evolution of communication. Animal communication and noise. Springer, Berlin, Heidelberg, pp. 7–30

Wiley RH, Richards DG (1978) Physical constraints on acoustic communication in the atmosphere: implications for the evolution of animal vocalizations. Behav Ecol Sociobiol 3 (1):69–94

Wong BM, Candolin U (2005) How is female mate choice affected by male competition? Biol Rev 80:559–571

Wyttenbach RA, Farris HE (2004) Psychophysics in insect hearing. Microsc Res Tech 63 (6):375–387

Yost WH (1994) Fundamentals of hearing: an introduction. Academic Press, San Diego, CA

Zuk M, Kolluru GR (1998) Exploitation of sexual signals by predators and parasitoids. Q Rev Biol 73:415–438

Zylinksi S, Johnsen S (2011) Mesopelagic cephalopods switch between transparency and pigmentation to optimize camouflage in the deep. Curr Biol 21:1937–1941

Chapter 7
Social Recognition in Anurans

Mark A. Bee

Abstract Learning to recognize and categorize other individuals is a cornerstone of animal social behavior. By learning about individually distinctive signal properties, receivers can perceptually discriminate among conspecifics to direct appropriate behaviors toward particular individuals. One context for social recognition arises from contests over territories. In many species, territory residents exhibit reduced levels of aggression toward nearby neighbors, but maintain a readiness to respond aggressively to unfamiliar individuals. Territory residents and their neighbors, which remain competitive rivals despite reaching a truce, are often described as "dear enemies." Although neighbor recognition is widespread across taxa, we have yet to satisfactorily elucidate the ecological and social factors that favor its evolution nor do we fully understand its underlying perceptual and cognitive mechanisms and how they potentially differ across species. Comparative and integrative studies of anurans (frogs and toads) have potential to address these gaps in current knowledge. After a brief introduction and primer on social recognition, this chapter critically reviews previous and ongoing work on vocally mediated neighbor recognition in territorial anurans. The focus is on comparing behavioral studies of recognition across species in light of similarities and differences in various ecological and social factors. Next, the chapter reviews studies aimed at elucidating the perceptual and cognitive mechanisms by which neighbor recognition is achieved by one particularly well-studied species. By adopting a case-study perspective, this chapter outlines the promise of comparative and integrative approaches to investigating the evolution and psychological mechanisms of social recognition in anurans, while also illustrating the perils that arise when inappropriate or inadequate methodologies are used in these investigations.

M.A. Bee (✉)
Department of Ecology, Evolution, and Behavior, Graduate Program in Neuroscience,
University of Minnesota – Twin Cities, St. Paul, MN, USA
e-mail: mbee@umn.edu

M.A. Bee, C.T. Miller (eds.), *Psychological Mechanisms in Animal Communication*, Animal Signals and Communication 5,
DOI 10.1007/978-3-319-48690-1_7

7.1 Introduction

Amphibians have had a reputation problem for centuries. As Linnaeus (1758) observed of “these most terrible and vile animals” in his *Systema Naturae*, “most amphibians are rough, with a cold body, a ghastly color, cartilaginous skeleton, foul skin, fierce face, a meditative gaze, a foul odor, a harsh call, a squalid habitat, and terrible venom. Their Author has not, therefore, done much boasting on their account” (translated by Kitchell and Dundee 1994). Although we may now consider them less terrible and vile than Linnaeus once did, the wonderfully interesting behaviors of amphibians remain underappreciated by the public and in certain scientific circles too. Perhaps part of the reason for this is that we humans, as an extremely social species, tend to see less of ourselves in amphibians than we see in other highly social animals like meerkats (Chap. 8) and monkeys (Chaps. 9 and 10). But it would be a serious mistake to presume that social behavior plays no part in the lives of amphibians (Wells 1977, 2007). Some anuran amphibians, for example, display rich social behaviors, including complex and coordinated biparental care of offspring and monogamous relationships (e.g., Brown et al. 2010; Tumulty et al. 2014). Male anurans are famous for the large social aggregations they form for the purpose of breeding. In these “choruses,” males compete to attract females by producing loud advertisement calls (Wells and Schwartz 2007). Females are often quite selective for high-quality males of their own species based on assessing the vocal behavior of potential mates (Ryan and Keddy-Hector 1992; Gerhardt and Huber 2002). Males directly compete against each other vocally using both advertisement calls and distinct aggressive calls, and many species engage in direct physical combat (Dyson et al. 2013; Bee et al. 2016). These aggressive interactions typically arise over access to females directly, access to calling sites free from the acoustic clutter generated by nearby males, or access to territories that contain breeding resources, such as oviposition sites or tadpole-rearing sites. Male anurans can make decisions about escalating contests by assessing their own or their opponent’s size and fighting ability (reviewed in Bee et al. 2016). As I will review in this chapter, the males of some species learn to recognize and display lower levels of aggression toward potentially frequent opponents.

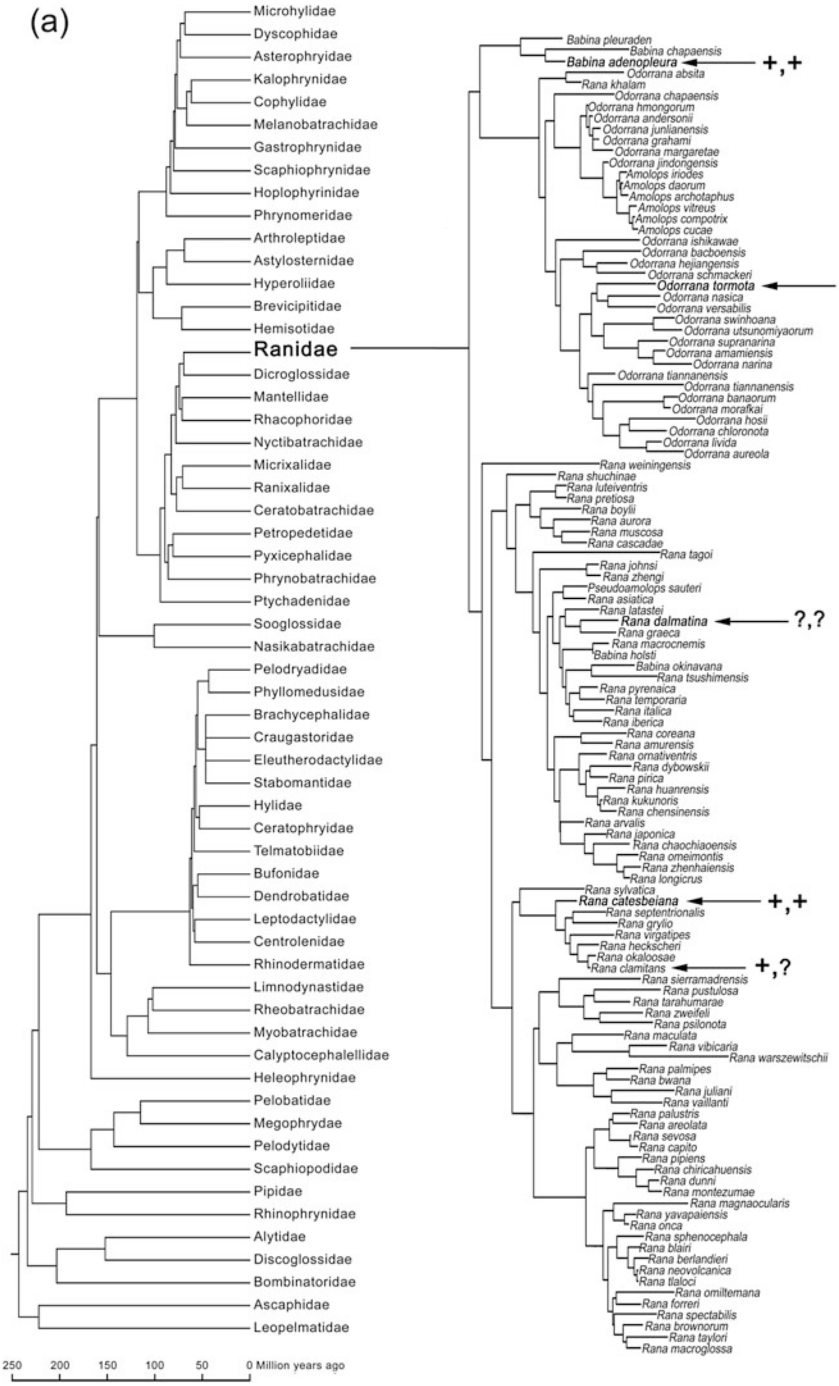
Anuran mating systems and social behavior are evolutionarily labile and highly dependent on the temporal and spatial distributions of resources required for breeding (reviewed in Wells 1977, 2007). Differences in resource availability select for different kinds of mating systems and social behaviors that fall along a continuum. At one end, “explosive breeders” have very short breeding periods lasting just one or a few days. Males often engage in scramble competition for access to females as they arrive at a breeding site. There is little defense of long-term calling sites or territories among explosive breeders. At the opposite end of this continuum are “prolonged breeders,” which, as the phrase suggests, have longer breeding seasons lasting weeks or months, not days. Prolonged breeders are more likely to defend calling sites or establish territories. When males defend an area that contains no breeding resources for females, but instead functions primarily as a calling site, the mating system is often described as lek polygyny (e.g., Bourne 1992). If the defended area contains breeding resources, and females oviposit in these areas, the mating system is instead more often described as resource defense polygyny (e.g., Howard 1978a, b). Individuals of some prolonged breeding species defend their territories for long periods of time, in some cases for weeks, months, and even years. Consequently, they may also share territorial boundaries with particular neighbors over similarly long periods.

In other territorial groups of animals, neighbors that share long-term territorial boundaries often display a so-called "dear enemy" effect (Fisher 1954). Neighbors that share a territorial boundary are considered dear enemies when they exhibit lower levels of aggression toward each other compared to the level of aggression each exhibits toward non-territorial individuals ("strangers") that intrude into their territory (Temeles 1994). A dear enemy effect implies neighbor recognition, but neighbor recognition can also occur in the absence of a dear enemy effect, for example, if neighbors are familiar but also as threatening as strangers (Temeles 1994). The form of social recognition evidenced by the dear enemy effect has been particularly well documented in acoustic playback experiments with songbirds (Falls 1982; Lambrechts and Dhondt 1995; Stoddard 1996). Not only do territorial songbirds perceptually and behaviorally discriminate between the songs of familiar neighbors and unfamiliar strangers, they also discriminate between the songs of different familiar neighbors and associate a familiar neighbor's songs with a particular direction corresponding to the location of the neighbor's territory. The abilities to vocally recognize territorial neighbors and associate them with specific locations engage a potentially complex suite of perceptual and cognitive mechanisms.

To what extent do territorial anurans demonstrate social recognition abilities similar to those of other territorial animals, such as songbirds, and what do we know about the mechanisms underlying such abilities? These are the two main questions addressed in this chapter. After a brief discussion of social recognition systems in Sect. 7.2, this chapter is divided into two main parts. In Sect. 7.3, I critically review previous studies that have investigated neighbor recognition and the dear enemy effect in anurans (see also Bee et al. 2016). Furthermore, I attempt to do so within the context of considering ecological and social factors associated with each species' breeding behavior. Neighbor recognition and the dear enemy effect have been investigated in only six frog species from just two families, the Ranidae (Fig. 7.1a) and Dendrobatidae (Fig. 7.1b).[1] As will become clear in Sect. 7.3, what we know of these few species is constrained by problems of methodology that impact interpretations of results. Given that there are more than 6600 recognized anuran species, our current state of knowledge on social recognition in anurans remains extremely limited and invites further study. In Sect. 7.4, I review

---

[1] There have been recent upheavals in anuran taxonomy, with debates about nomenclature still ongoing (Frost et al. 2006; Pyron and Wiens 2011; Frost 2015; AmphibiaWeb 2016). In this chapter, my use of species names and higher order taxonomy follows AmphibiaWeb (2016).


(a)
Microhylidae
Dyscophidae
Asterophryidae
Kalophrynidae
Cophylidae
Melanobatrachidae
Gastrophrynidae
Scaphiophrynidae
Hoplophyrinidae
Phrynomeridae
Arthroleptidae
Astylosternidae
Hyperoliidae
Brevicipitidae
Hemisotidae
Ranidae
Dicroglossidae
Mantellidae
Rhacophoridae
Nyctibatrachidae
Micrixalidae
Ranixalidae
Ceratobatrachidae
Petropedetidae
Pyxicephalidae
Phrynobatrachidae
Ptychadenidae
Sooglossidae
Nasikabatrachidae
Pelodryadidae
Phyllomedusidae
Brachycephalidae
Craugastoridae
Eleutherodactylidae
Stabomantidae
Hylidae
Ceratophryidae
Telmatobiidae
Bufonidae
Dendrobatidae
Leptodactylidae
Centrolenidae
Rhinodermatidae
Limnodynastidae
Rheobatrachidae
Myobatrachidae
Calyptocephalellidae
Heleophrynidae
Pelobatidae
Megophrydae
Pelodytidae
Scaphiopodidae
Pipidae
Rhinophrynidae
Alytidae
Discoglossidae
Bombinatoridae
Ascaphidae
Leopelmatidae
250
200
150
100
50
0 Million years ago
Babina pleuraden
Babina chapaensis
Babina adenopleura
+,+
Odorrana absita
Rana khalam
Odorrana chapaensis
Odorrana hmongorum
Odorrana andersonii
Odorrana junlianensis
Odorrana grahami
Odorrana margaretae
Odorrana jindongensis
Amolops iriodes
Amolops daorum
Amolops archotaphus
Amolops vitreus
Amolops compotrix
Amolops cucae
Odorrana ishikawae
Odorrana bacboensis
Odorrana hejiangensis
Odorrana schmackeri
Odorrana tormota
+,+
Odorrana nasica
Odorrana versabilis
Odorrana swinhoana
Odorrana utsunomiyaorum
Odorrana supranarina
Odorrana amamiensis
Odorrana narina
Odorrana tiannanensis
Odorrana tiannanensis
Odorrana banaorum
Odorrana morafkai
Odorrana hosii
Odorrana chloronota
Odorrana livida
Odorrana aureola
Rana weiningensis
Rana shuchinae
Rana luteiventris
Rana pretiosa
Rana boylii
Rana aurora
Rana muscosa
Rana cascadae
Rana tagoi
Rana johnsi
Rana zhengi
Pseudoamolops sauteri
Rana asiatica
Rana latastei
Rana dalmatina
?,?
Rana graeca
Rana macrocnemis
Babina holsti
Babina okinavana
Rana tsushimensis
Rana pyrenaica
Rana temporaria
Rana italica
Rana iberica
Rana coreana
Rana amurensis
Rana ornativentris
Rana dybowskii
Rana pirica
Rana huanrensis
Rana kukunoris
Rana chensinensis
Rana arvalis
Rana japonica
Rana chaochiaoensis
Rana omeimontis
Rana zhenhaiensis
Rana longicrus
Rana sylvatica
Rana catesbeiana
+,+
Rana septentrionalis
Rana grylio
Rana virgatipes
Rana heckscheri
Rana okaloosae
Rana clamitans
+,?
Rana sierramadrensis
Rana pustulosa
Rana tarahumarae
Rana zweifeli
Rana psilonota
Rana maculata
Rana vibicaria
Rana warszewitschii
Rana palmipes
Rana bwana
Rana juliani
Rana vaillanti
Rana palustris
Rana areolata
Rana sevosa
Rana capito
Rana pipiens
Rana chiricahuensis
Rana dunni
Rana montezumae
Rana magnaocularis
Rana yavapaiensis
Rana onca
Rana sphenocephala
Rana blairi
Rana berlandieri
Rana neovolcanica
Rana tlaloci
Rana omiltemana
Rana forreri
Rana spectabilis
Rana brownorum
Rana taylori
Rana macroglossa


(b)
Microhylidae
Dyscophidae
Asterophryidae
Kalophrynidae
Cophylidae
Melanobatrachidae
Gastrophrynidae
Scaphiophrynidae
Hoplophyrinidae
Phrynomeridae
Arthroleptidae
Astylosternidae
Hyperoliidae
Brevicipitidae
Hemisotidae
Ranidae
Dicroglossidae
Mantellidae
Rhacophoridae
Nyctibatrachidae
Micrixalidae
Ranixalidae
Ceratobatrachidae
Petropedetidae
Pyxicephalidae
Phrynobatrachidae
Ptychadenidae
Sooglossidae
Nasikabatrachidae
Pelodryadidae
Phyllomedusidae
Brachycephalidae
Craugastoridae
Eleutherodactylidae
Stabomantidae
Hylidae
Ceratophryidae
Telmatobiidae
Bufonidae
Dendrobatidae
Leptodactylidae
Centrolenidae
Rhinodermatidae
Limnodynastidae
Rheobatrachidae
Myobatrachidae
Calyptocephalellidae
Heleophrynidae
Pelobatidae
Megophrydae
Pelodytidae
Scaphiopodidae
Pipidae
Rhinophrynidae
Alytidae
Discoglossidae
Bombinatoridae
Ascaphidae
Leopelmatidae
250 200 150 100 50 0 Million years ago
Rheobates palmatus
Anomaloglossus stepheni
Anomaloglossus baeobatrachus
Anomaloglossus tepuyensis
Anomaloglossus degranvillei
Anomaloglossus praderioi
Anomaloglossus roraima
Anomaloglossus beebei
+,+
Aromobates saltuensis
Aromobates nocturnus
Mannophryne herminae
Mannophryne trinitatis
Mannophryne venezuelensis
Allobates undulatus
Allobates talamancae
Allobates brunneus
Allobates nidicola
Allobates zaparo
Allobates femoralis
+,?
Allobates juanii
Allobates gasconi
Allobates caeruleodactylus
Allobates trilineatus
Allobates conspicuus
Silverstoneia nubicola
Silvertoneia flotator
Epipedobates boulengeri
Epipedobates espinosai
Epipedobates machalilla
Epipedobates anthonyi
Epipedobates tricolor
Colostethus latinasus
Colostethus pratti
Colostethus imbricolus
Colostethus inguinalis
Colostethus panamansis
Colostethus fraterdanieli
Colostethus fugax
Colostethus argyrogaster
Ameerega altamazonica
Ameerega simulans
Ameerega silverstonei
Ameerega flavopicta
Ameerega braccata
Ameerega bassleri
Ameerega bilinguis
Ameerega parvula
Ameerega pongoensis
Ameerega trivittata
Ameerega pulchripecta
Ameerega hahneli
Ameerega picta
Ameerega macero
Ameerega rubriventris
Ameerega smaragdina
Ameerega cainarachi
Ameerega petersi
Hyloxalus subpunctatus
Hyloxalus leucophaeus
Hyloxalus sordidatus
Hyloxalus maculosus
Hyloxalus bocagei
Hyloxalus sauli
Hyloxalus vertebralis
Hyloxalus pulchellus
Hyloxalus delatorreae
Hyloxalus anthracinus
Hyloxalus pulcherrimus
Hyloxalus sylvaticus
Hyloxalus nexipus
Hyloxalus azureiventris
Hyloxalus chlorocraspedus
Hyloxalus idiomelus
Hyloxalus insulatus
Hyloxalus elachyhistus
Hyloxalus infraguttatus
Hyloxalus toachi
Hyloxalus awa
Phyllobates vittatus
Phyllobates lugubris
Phyllobates bicolor
Phyllobates aurotaenia
Phyllobates terribilis
Minyobates steyermarki
Adelphobates quinquevittatus
Adelphobates castaneoticus
Adelphobates galactonotus
Dendrobates truncatus
Dendrobates auratus
Dendrobates tinctorius
Dendrobates leucomelas
Oophaga granulifera
Oophaga histrionica
Oophaga sylvatica
Oophaga lehmanni
Oophaga arborea
Oophaga speciosa
Oophaga vicentei
Oophaga pumilio
+,-
Excidobates captivus
Excidobates mysteriosus
Andinobates virolinensis
Andinobates bombetes
Andinobates fulguritus
Andinobates minutus
Andinobates claudiae
Ranitomeya sirensis
Ranitomeya sirensis
Ranitomeya flavovittata
Ranitomeya imitator
Ranitomeya vansolinii
Ranitomeya fantastica
Ranitomeya ventrimaculata
Ranitomeya reticulata
Ranitomeya uakarii
Ranitomeya amazonica
Ranitomeya variabilis
Ranitomeya ventrimaculata

work on one particularly well-studied species, the North American bullfrog, *Rana catesbeiana* (Ranidae), that has investigated the acoustic, perceptual, and cognitive bases of neighbor recognition in some detail.

7.2 Components of Social Recognition Systems

The concept and importance of social recognition systems in animals have been reviewed in detail elsewhere (Beecher 1982, 1988, 1989a, b, 1991; Colgan 1983; Halpin 1986; Beecher et al. 1989; Sherman et al. 1997; Starks 2004; Bee 2006; Tibbetts and Dale 2007; Bradbury and Vehrencamp 2011; Wiley 2013). For brevity's sake, I will only mention here several key components of social recognition systems that pertain to neighbor recognition in anurans. I do not want to get bogged down too much with definitions. Earlier attempts in the literature to draw stark contrasts between "individual discrimination" and "individual recognition," or to define "true individual recognition" (as if there were such a thing as "false individual recognition"), highlight the potential futility of doing so. (See Chap. 8 for an alternative view.) Rather, for our purposes, I will consider social recognition operationally to mean *familiarity with one or more distinctive features of the signals of one or more conspecific individuals that is acquired through learning and that facilitates functional behavioral discrimination between two or more social categories of individuals.* Although this operational definition emphasizes the role of learning about identity signals, it could be broadened easily to accommodate recognition that is unlearned or based on other phenotypic traits or genotypes. Four elements of this definition merit brief elaboration.

First, a key component of many social recognition systems is the ability to perceptually discriminate among different categories of individuals. Hence, cogni-

Fig. 7.1 Anuran social recognition in a phylogenetic framework. Shown on the *left* in both panel (**a**) and panel (**b**) is the same "timetree" of anuran families (redrawn from Bossuyt and Roelants 2009). Social recognition in the form of the dear enemy effect has so far only been investigated in (**a**) the Ranidae and (**b**) the Dendrobatidae. Shown on the *right* in both panel (**a**) and panel (**b**) are portions of a large-scale maximum likelihood estimate of amphibian phylogeny (redrawn from Pyron and Wiens 2011) illustrating the hypothesized evolutionary relationships between anuran species in which the dear enemy effect has been investigated. The two symbols next to a species' name in (**a**) and (**b**) indicate, respectively, whether the species has been shown to have individually distinctive calls (*left*: "+" yes, "–" no) and whether the species has been shown to exhibit the dear enemy effect (*right*: "+" yes, "–" no, "?" equivocal or untested). Data are shown for olive frogs (*Babina adenopleura*) (Chuang et al. 2017), concave-eared torrent frogs (*Odorrana tormota*) (Feng et al. 2009a, 2009b), agile frogs (*Rana dalmatina*) (Lesbarrères and Lodé 2002), bullfrogs (*Rana catesbeiana*) (Davis 1987; Bee and Gerhardt 2001b), green frogs (*Rana clamitans*) (Owen and Perrill 1998; Bee and Schachtman 2000; Bee et al. 2001), golden rocket frogs (*Anomaloglossus beebei*) (Bourne et al. 2001; Pettitt et al. 2013), brilliant-thighed poison frogs (*Allobates femoralis*) (Gasser et al. 2009), and strawberry poison frogs (*Oophaga pumilio*) (Bee 2003b; Pröhl 2003) tive mechanisms related to category formation and categorization are key (Bee 2006; Miller and Bee 2012). Categories might be very broad, such as familiar versus unfamiliar or kin versus non-kin. Alternatively, they might be so narrow as to correspond to single individuals, such that each category corresponds to a different social partner. Recognition of just one individual, such as a mate, automatically creates two categories (e.g., mate versus non-mate). Animals might also discriminate among more than two categories. Wiley (2013) has described these issues in terms of the "specificity" (how many individuals per category) and "multiplicity" (how many categories) of recognition systems. For our purposes, I will generally consider discrimination between the categories of "neighbor" and "stranger" (or between "familiar" and "unfamiliar"). As shorthand, I will call this "neighbor recognition." Most studies of territorial neighbor recognition in anurans have presented subjects with the calls of a single neighbor (or a presumably familiar individual) and a stranger from the direction of the neighbor's territory. I know of no definitive study of anurans that has shown they can learn to recognize multiple different neighbors.

Second, signals have to be distinctive among the to-be-discriminated categories. In this review, I discuss this issue in terms of "individual vocal distinctiveness," since categorizing one animal as a neighbor versus a stranger requires that their signals be individually distinct, as we would generally not expect a group signature to exist for the category of "neighbors" (but see Briefer et al. 2008). Individually distinct signals, also known as identity signals, are those with properties that reliably vary more among individuals than within individuals. Manser (Chap. 8) incorporates information on individual identity into a broader framework for considering referents in animal vocalizations. There are a variety of ways to characterize individual distinctiveness based on acoustic recordings (reviewed in Bee et al. 2016). After analyzing the acoustic properties of vocalizations, for example, some authors describe the among-individual and within-individual variation using coefficients of variation (CV). Individually distinctive properties are those having a ratio of among-individual variation (CV_a) to within-individual variation (CV_w) greater than 1.0. Model II ANOVA can be used to statistically compare among-individual variation to within-individual variation, though caution is needed when interpreting outcomes of these analyses when signal properties are intercorrelated (Beecher 1989b). Estimates of "repeatability," a concept borrowed from quantitative genetics (Lessells and Boag 1987; Boake 1989), can be derived from ANOVA and used to quantify the relative magnitude of among-individual and within-individual variation. Discriminant function analysis (DFA) is another multivariate statistical approach used to quantify individual distinctiveness in signals. DFA can be used to statistically assign, or classify, signals to the individuals that produced them. Individually distinctive signals yield a classification success significantly higher than expected by chance. Beecher (1989b) has also described methods for quantifying the information content of signals as it pertains to individual identity. Acoustic and statistical analyses of signals, of course, can only characterize the *potential* for behavioral discrimination between social categories.

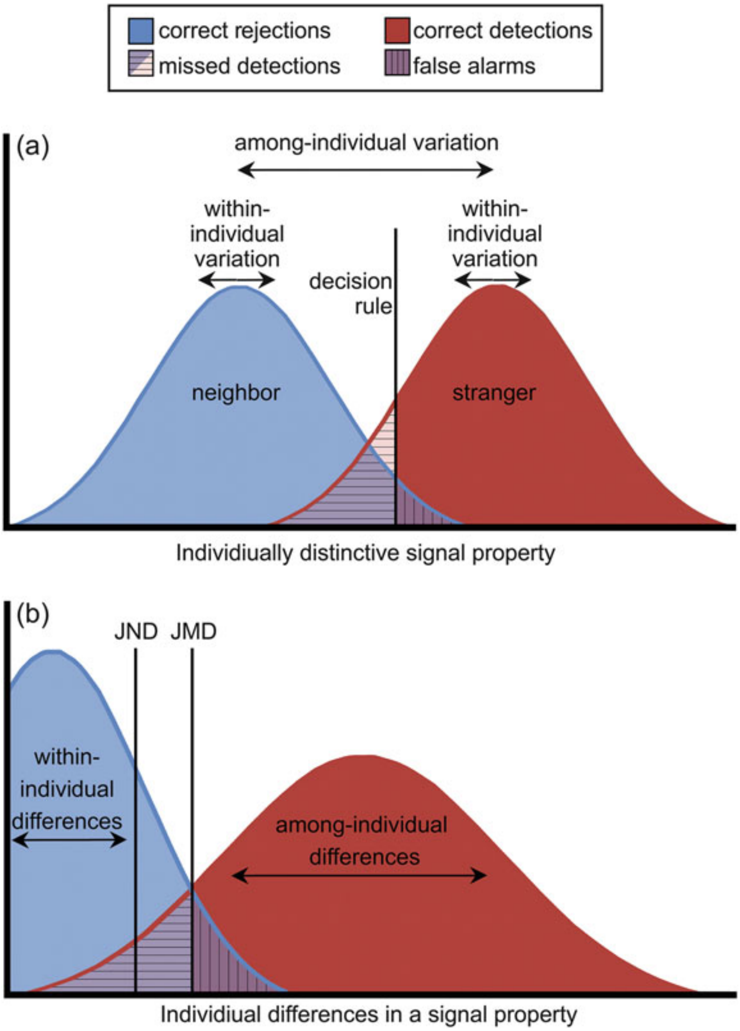


Fig. 7.2 A signal detection theory (SDT) view of the problem of detecting a stranger that represents a threat to territory ownership. In this scenario, a signaling neighbor represents the "noise" in which the signals of a stranger must be detected. Two types of correct decisions are possible. A "correct detection" occurs when the territorial resident correctly identifies and responds to a perceived signal as a threat because it is the signal of a stranger. A "correct rejection" occurs when the territorial resident correctly identifies the signal of a neighbor and withholds an aggressive response. Two types of errors are also possible. A "missed detection" occurs when the territorial resident incorrectly decides that the signal of a stranger was produced by a neighbor and withholds an aggressive response. A "false alarm" occurs when the territorial resident incorrectly decides that a signal of a neighbor was produced by a stranger and responds aggressively by attacking its neighbor. Evolution should optimally place a receiver's decision rule depending on the relative benefits and costs of both correct and incorrect decisions, as well as the likelihood of encountering different types of signalers (i.e., neighbors versus strangers). In panel (**a**), the individually distinctive signals of a neighbor and stranger are represented as probability density

Empirical evidence from behavioral studies is always necessary to conclude that social recognition occurs.

Third, from a receiver's perspective, determining whether or not a perceived signal was produced by a neighbor or stranger amounts to a signal detection problem. From the perspective of *signal detection theory* (*SDT*) (Bradbury and Vehrencamp 2011; Wiley 2015), the signals of a neighbor represent the "noise" against which the signal of a stranger must be detected (Fig. 7.2). The among-individual differences between the signals generated by a neighbor and a potential stranger are what make signal detection possible. But the signals of both neighbors and strangers also vary within individuals. This within-individual variability is one important source of noise in communication-based social recognition. Receivers must have some kind of threshold, or decision rule, to determine whether a perceived signal was produced by a neighbor or a stranger. When applying a decision rule, errors are inevitable (Fig. 7.2; Wiley 2015). The precise location of a receiver's decision rule depends on the benefits of correct decisions and the costs of wrong decisions, as well as on the likelihood of encountering neighbors versus strangers. Decision rules are generally expected to be evolutionarily labile and perhaps even plastic within individuals (e.g., Stoddard 1996) depending on the timescales of variation in the relative threats posed by neighbors and strangers.

Finally, because recognition involves perceptual and cognitive processes that are inherently internal to the receiver, empirical evidence for recognition must be based on overt behavioral discrimination by receivers between signals from two or more social categories. Field studies of social recognition in animals typically exploit aspects of the animal's natural behavioral repertoire in response to natural stimuli (Chaps. 8 and 9). In such studies, a lack of behavioral discrimination might result for at least two different reasons (Chap. 6). Differences between the signals

◄

Fig. 7.2 (continued) functions along an x-axis that corresponds to a single individually distinctive signal property (e.g., fundamental frequency). Signal properties vary both within individuals (*short horizontal arrows*) and among individuals (*long horizontal arrow*). Receivers must employ a decision rule (*vertical line*) for deciding whether or not a perceived signal matches their memory of a neighbor's signals (see also Chap. 6). In this example, receivers respond aggressively toward the signal producer whenever they perceive a signal value that falls to the right of the decision rule along the x-axis. Overlap in signal properties and the placement of the receiver's decision rule create the four possible outcomes according to this SDT scenario. Panel (**b**) depicts another way to conceptualize the problem in the form of probability density functions that correspond to *differences* in signals. Here, one distribution corresponds to the differences in a particular signal property that occur between repeated renditions of the signal by the same individual (within-individual differences). These differences reflect within-individual variation. A second distribution corresponds to the differences in the same signal property that occur among the individuals in the population and reflect among-individual variation. Here, the receiver's decision rule for behavioral discrimination corresponds to a "just-meaningful difference" (JMD) that has been placed where the magnitude of among-individual differences first exceed within-individual differences. Note that receivers may be able to perceptually discriminate smaller differences that do not elicit discriminative behavioral responses. This is indicated by the "just-noticeable difference" (JND) to the left of the JMD generated by different social categories may be smaller than the resolving power of the animal's sensory and perceptual systems, that is, smaller than the animal's *just-noticeable difference* (*JND*). In this case, there would be a failure of both perceptual and behavioral discrimination. However, differences between the signals of different social categories may well exceed an animal's JND but not be large enough to be meaningful to the animal and, therefore, may not exceed what Nelson and Marler (1990) have called the animal's *just-meaningful difference* (*JMD*). In these cases, perceptual discrimination can occur in the absence of behavioral discrimination. Thus, researchers must always bear in mind that a negative result showing that animals *do not* behaviorally discriminate between two social categories does not necessarily mean the animals *cannot* perceptually discriminate between them (Chap. 6). In the context of the dear enemy effect, the expected behavioral discrimination is that subjects respond more aggressively toward strangers than toward their neighbors. Evidence for discrimination could involve differences in vocal output in response to neighbors and strangers, with relatively more aggressive signals given in response to strangers, or a greater tendency of territory residents to display toward, approach and make physical contact with strangers or speakers broadcasting their vocalizations compared with real or acoustically simulated neighbors. Aggressive signaling and other aggressive behaviors are well described in many anurans (Dyson et al. 2013; Bee et al. 2016), so most species probably possess some potential to flexibly display or withhold aggressive behaviors.

7.3 Neighbor Recognition and the Dear Enemy Effect in Anurans

In this section, I critically review, on a case-by-case basis, previous studies of vocally mediated neighbor recognition and the dear enemy effect in anurans. For each species, I first set the stage by briefly describing key features of the animal's natural history and behavior. The goal is to provide some necessary context about ecological and social factors related to the species' breeding behavior. This is followed by a review of empirical playback studies of behavioral discrimination between neighbors and strangers based on individual differences in advertisement calls (Fig. 7.3). All of the species described in this section have been reported to have individually distinctive vocalizations based on descriptive studies that have analyzed patterns of individual variation in advertisement calls (reviewed in Bee et al. 2016). I begin by reviewing research on one of the most well-studied and iconic of all North American anurans, the bullfrog.

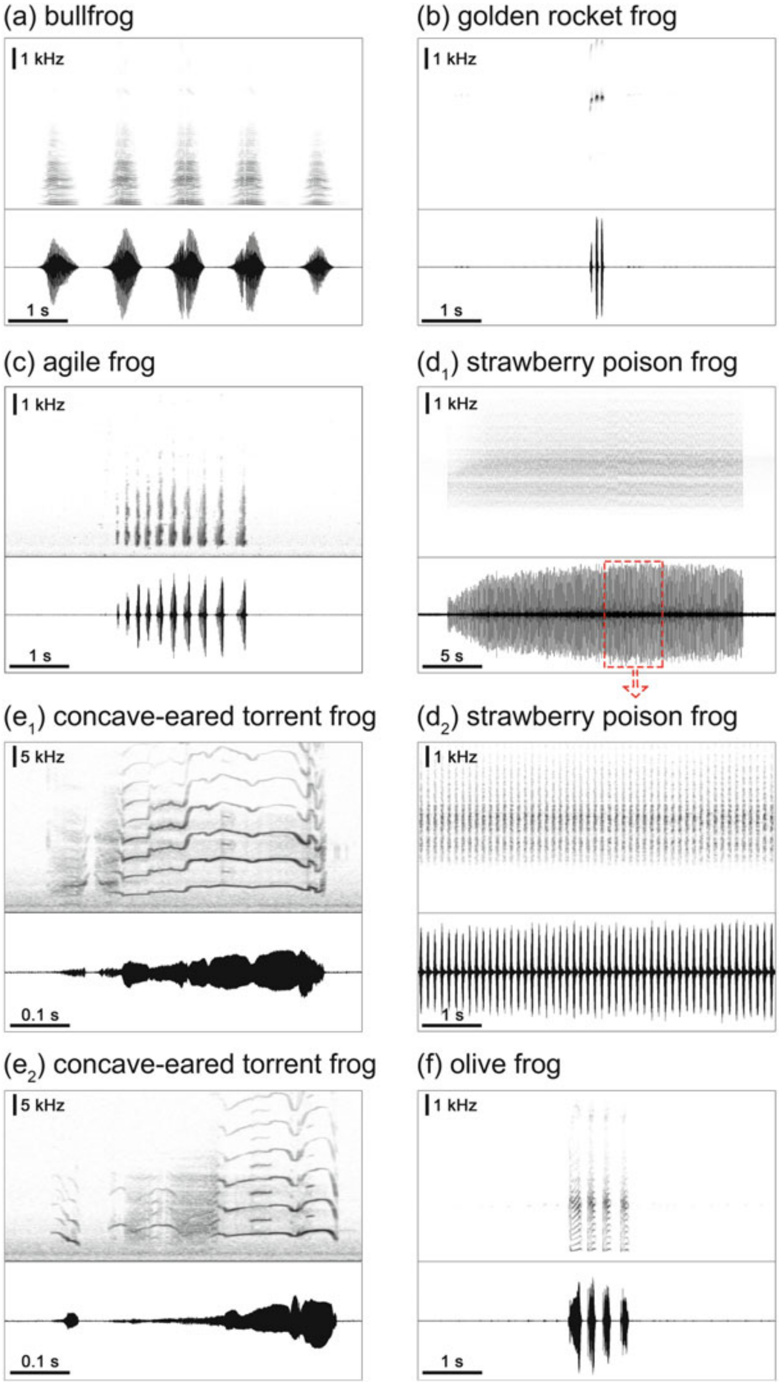


Fig. 7.3 Spectrograms (*top traces*) and waveforms (*bottom traces*) of the advertisement calls produced by the males of the species discussed in this chapter (FFT = 1024 or 2048 points). *Horizontal scale bars* denote time (in s) and *vertical scale bars* denote frequency (in kHz).

7.3.1 Bullfrogs (Ranidae)

7.3.1.1 Bullfrogs: Natural History

It was not long after the publications of Wilson's (1975) *Sociobiology: The New Synthesis*, marking the dawn of behavioral ecology, and Wells' (1977) seminal paper on anuran social and reproductive behaviors that North American bullfrogs (*Rana catesbeiana*) claimed center stage in studies describing anuran mating and territorial behaviors and quantifying factors that influence reproductive success (Emlen 1976; Howard 1978a, b, 1979, 1981, 1983, 1984; Ryan 1980). These studies reported on key features of bullfrog behavior that would lay a solid foundation for future studies of neighbor recognition in this species. The species' mating system is characterized as resource defense polygyny, though as predicted by Wells (1977), there is some variation in the mating system, with males apparently engaging in lek polygyny prior to sorting out stable territorial boundaries with their neighbors. Males establish and defend territories on the surface of the water or along the bank of ponds, lakes, and streams. Females oviposit in these territories, in which the structure of submerged vegetation turns out to be an important breeding resource. Not all territories are created equally, and these early studies revealed that larger males win more fights over territories and they possess better territories, as measured, for example, in terms of embryo survival. Individual males can occupy the same territory for more than 4 weeks, and adjacent neighbors can share a common territorial boundary for more than 2 weeks (Haas 1977; Bee 2001b). While floating in its territory, a male bullfrog produces advertisement calls consisting of one or more (usually about six or seven) long notes or "croaks," each lasting about 700 ms and separated from the next by about 500 ms (Fig. 7.3a). Studies investigating patterns of individual distinctiveness in bullfrog advertisement calls are reviewed in Sect. 7.4.1. When approached by an intruder who calls in close proximity, a territorial male will use acoustically distinctive aggressive calls (Wiewandt 1969) and approach the intruder in a splashy display. If the intruder persists, territorial contests can escalate to all-out physical combat (Fig. 7.4). Larger males typically win more contests, and the most escalated contests occur between males of similar size. Of greatest relevance to considering vocally mediated social recognition in bullfrogs are the facts that males are territorial and aggressive and exhibit reasonably long-term defense of a breeding resource.

Fig. 7.3 (continued) (**a**) Five-note call from a North American bullfrog (*Rana catesbeiana*). (**b**) Three-note call from a golden rocket frog (*Anomaloglossus beebei*). (**c**) Pulsatile call of an agile frog (*Rana dalmatina*); recording courtesy of Vanessa Sarasola, Rafael Marquez, and www.Fonozoo.com. (**d**) Call group of a strawberry poison frog (*Oophaga pumilio*) depicted at two different timescales ($\mathbf{d_1}$ and $\mathbf{d_2}$); the *highlighted box* in $\mathbf{d_1}$ shows the portion of the call group depicted in $\mathbf{d_2}$. (**e**) Two examples ($\mathbf{e_1}$ and $\mathbf{e_2}$) of a "long call" produced by two different males of the concave-eared torrent frog (*Odorrana tormota*); recordings courtesy of Albert Feng, Peter Narins, and Junxian Shen. (**f**) Four-note call from an olive frog (*Babina adenopleura*)

Fig. 7.4 Two male bullfrogs (*Rana catesbeiana*) fighting over a territory. Shown here is a series of photographs (**a**–**d**) of a fight sequence. (**a**) Two males initially face off while directing advertisement calls and distinctive aggressive calls toward their opponent. (**b**) The encounter eventually escalates to physical combat in which the two males wrestle and grapple with each other. (**c**, **d**) One male (the eventual loser) is flipped onto its back and submerged below the surface of the water. These sorts of fights are common, especially early during the breeding season, as neighboring males sort out stable territory boundaries. Photos copyright Robert McCaw (www.robertmccaw.com), used with permission

7.3.1.2 Bullfrogs: Neighbor Recognition

Davis (1987) investigated whether territorial male bullfrogs could discriminate between their neighbors and strangers based on individual differences in advertisement calls. In the first of two experiments, he used a within-subjects design to present territorial males with the prerecorded calls of an adjacent territorial neighbor or an unfamiliar stranger. After removing the neighbor, sounds were presented from a speaker floating on a Styrofoam platform positioned in the neighbor's territory. Playbacks of the two stimulus types were separated by brief time-outs of 15–20 min, and stimulus order was counterbalanced across subjects. Davis (1987) counted the numbers of advertisement calls and aggressive calls elicited by each stimulus, and he measured the number of meters subjects approached toward the playback speaker. Ten of the 11 subjects responded more strongly to playbacks of the calls of a stranger compared with those of their neighbor (Fig. 7.5a). No males exhibited more escalated responses to the calls of their neighbor compared with those of a stranger. These results, which were consistent

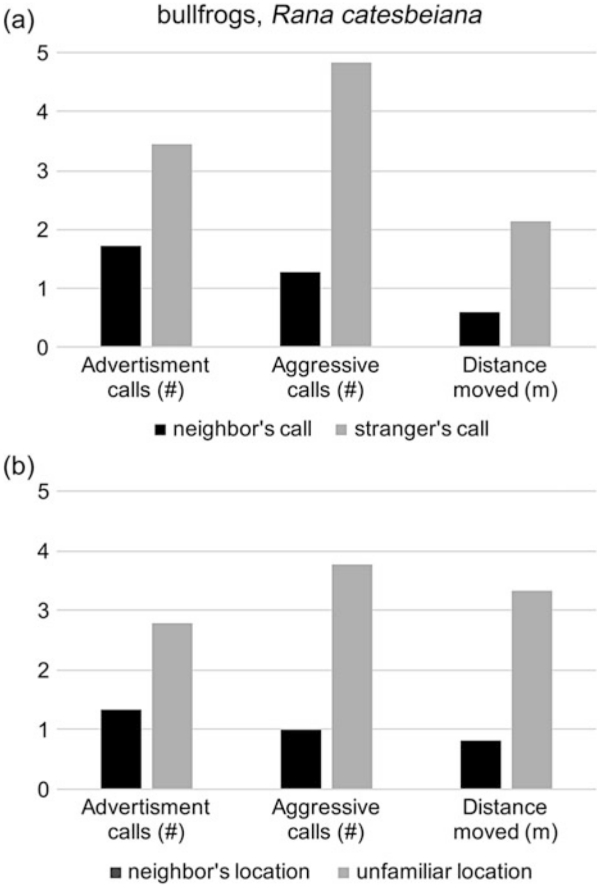


Fig. 7.5 Results from a playback study of the dear enemy effect in bullfrogs (*Rana catesbeiana*). Results shown are redrawn from Davis (1987) and depict the mean numbers of advertisement calls and aggressive calls produced and the distances moved toward the playback speaker in response to (**a**) the calls of a neighbor or stranger broadcast from the direction of the neighbor's usual territory and (**b**) the calls of a neighbor broadcast from the direction of the neighbor's usual territory or from a new, unfamiliar location on the opposite side of the subject's territory

with numerous previous studies of neighbor recognition in territorial songbirds, were the first data (and for nearly 15 years, the only data) indicating that territorial frogs can perceptually and behaviorally discriminate between the calls of neighbors and strangers in ways consistent with the dear enemy effect.

But Davis (1987) went one step further in a second experiment, in which he presented the advertisement calls of each subject's neighbor from both the neighbor's usual territorial position and from the opposite side of the subject's territory. Similar designs had been used in previous studies of songbirds to show that residents associate the songs of neighbors with particular locations. The typical outcome of such experiments in songbirds was that subjects responded more aggressively toward the familiar songs of neighbors presented from an incorrect or novel location. And this is what Davis (1987) found too. All nine of his subjects responded more aggressively toward a neighbor's call coming from an unfamiliar location compared to when it came from the neighbor's normal location (Fig. 7.5b).

The standard interpretation of such results is that territory holders can do more than merely discriminate between familiar and unfamiliar sounds (Beer 1970; Falls 1982). Rather, they can also learn to associate familiar sounds with a particular location.

That Davis's (1987) results with bullfrogs exhibited such remarkable similarity to those reported previously in many songbirds suggested some anurans might possess perceptual and cognitive abilities beyond those historically attributed to this group, for which learning in the context of communication is often dismissed. More to the point, Davis's (1987) results were important because they showed evolutionary convergence in the *behaviors* of territorial anurans and songbirds in the context of social recognition. To what extent there has been convergence in the perceptual and cognitive mechanisms underlying these behavioral similarities is a largely open question.

A number of methodological issues deserve discussion because they potentially qualify Davis's (1987) interpretation of his results. Two issues relate to the number of strangers used as stimulus donors and what they sounded like. First, though he typically describes his first experiment using the plural form "strangers," Davis's (1987) text is not entirely clear on whether his playbacks used a single stranger stimulus, a small subset of stranger stimuli, or a unique stranger stimulus for each neighbor. This uncertainty goes to the generalizability of his results, and some would say to the issue of pseudoreplication, though to be fair to Davis (1987), these contentious issues of playback design would not erupt in the literature for 2 more years (Catchpole 1989; Kroodsma 1989a, b, 1990; Searcy 1989; McGregor et al. 1992; Kroodsma et al. 2001). Second, Davis (1987) provided no description of the acoustic properties of his neighbor and stranger stimuli. Is there some chance that his stranger stimuli were (or stimulus was) perceived as somehow more threatening than neighbors, for example, by virtue of having acoustic properties that conveyed information about larger size, superior fighting ability, or higher motivation to fight? This seems unlikely given the results of later work showing that territorial bullfrogs did not behaviorally discriminate between stimuli having different fundamental frequencies, an acoustic property strongly predictive of body size, which in turn predicts fighting ability in this species (Bee 2002). Nevertheless, this uncertainty underscores a persistent need to carefully characterize the acoustic properties of sounds used as neighbor and stranger stimuli. A third issue concerns the presence of neighbors in the unfamiliar locations used in Experiment 2. Davis (1987) only reports that "If a second neighbor occupied a territory opposite the subject male..., then it was also removed prior to the playback session." This issue is relevant because it remains ambiguous as to whether subjects responded *only* to a change in a familiar neighbor's position or whether they responded to replacing a neighbor on one side of their territory with a neighbor from the opposite side. The latter scenario has been interpreted as evidence of recognition of multiple individuals in some songbird studies. Whether frogs can recognize multiple familiar neighbors thus remains an open and important question.

7.3.2 Golden Rocket Frogs (Dendrobatidae)

7.3.2.1 Golden Rocket Frogs: Natural History

Endemic to a small geographic range in the Kaieteur National Park of Guyana, the golden rocket frog, *Anomaloglossus* (formerly *Colostethus*) *beebei*, represents an interesting candidate for studies of neighbor recognition. The reproductive and social behaviors of this small, diurnal species were first described in detail by Bourne et al. (2001). My lab has studied its communication system since 2007 (Pettitt 2012; Pettitt et al. 2012, 2013). Individuals spend nearly their entire lives on giant tank bromeliads (*Brocchinia micrantha*). Males establish and defend small territories on the leaves of these plants that contain one or more phytotelmata, small wells at the base of leaves that collect rainwater and mist from the nearby Kaieteur Falls. Individual males have been observed to occupy the same territory in different years (J. Tumulty and M. A. Bee, unpublished data). Neighboring territorial males are often in close proximity, with distances between the centers of adjacent territories commonly between 2 and 3 m (J. Tumulty and M. A. Bee, unpublished data). Reproduction occurs on the plant, and females oviposit into a phytotelm defended by her mate. Thus, as in bullfrogs, males defend territories that contain important breeding resources for females in the form of oviposition sites. Both parents provide care to their offspring. Males spend time near phytotelmata guarding fertilized eggs and transport their tadpoles to alternative phytotelmata after hatching (Bourne et al. 2001; Pettitt 2012).

During the breeding season, about 40 % of a male's activity budget is spent vocalizing (Bourne et al. 2001). Nearby neighboring males frequently engage in antiphonal calling, in some ways similar to the marmoset vocal behavior described by Toarmino et al. (Chap. 10). Indeed, Bourne et al. (2001) describe calling as a "socially facilitated activity." Pettitt et al. (2012) published the first quantitative description of the vocal repertoire of this species. From within their territories, males use long-range advertisement calls to attract females (Fig. 7.3b), and close-range courtship calls to convince them to mate once they arrive. Advertisement calls consist of a short series of three or four high-pitched, rapidly repeated pulses (~35 ms in duration) produced at rates of 10–12 pulses/s (Fig. 7.3b). The dominant frequency of a pulse is the second harmonic and typically falls in the range of 4–6 kHz. Males produce an acoustically distinct aggressive call in response to territorial intrusions by other males (Bourne et al. 2001; Pettitt et al. 2012). Territorial males, therefore, possess the behavioral repertoire necessary to exhibit potentially higher levels of aggression toward strangers compared with neighbors. Pettitt et al. (2013) conducted detailed acoustic and statistical analyses of a large number of advertisement calls to identify the potential cues used to discriminate among individuals. Even after correcting for temperature-induced variation in calls, temporal properties related to the pulsatile structure of the call (pulse duration, inter-pulse interval, pulse rate) contributed most toward statistically discriminating among individuals. Dominant frequency, which incidentally was not correlated with size, was also an important contributor toward individual distinctiveness. Features related to the overall structure of the call (call duration, number of pulses) also contributed to individual distinctiveness, but less so than other properties.

7.3.2.2 Golden Rocket Frogs: Neighbor Recognition

Bourne et al. (2001) reported results from a preliminary field playback experiment to assess whether males exhibit relatively lower levels of aggression toward their nearby neighbors compared with strangers. The behavior of 25 territorial males in response to the calls of neighbors and non-neighbors was assessed using a two-choice phonotaxis test. Speakers were suspended by twine above a bromeliad leaf from a framework made of saplings, and calls were broadcast at naturalistic sound pressure levels. A "choice" of neighbor versus non-neighbor was scored when the male approached to within 15 mm of one of the two speakers. Males were significantly more likely to choose (i.e., approach) the speaker broadcasting non-neighbor calls, and they did so while emitting aggressive calls. In contrast, they called antiphonally in response to playbacks of a neighbor's calls or ignored them altogether.

The results of Bourne et al. (2001) provide tantalizing evidence to suggest males of this territorial species exhibit relatively lower levels of aggression toward neighbors than toward strangers. However, insufficient experimental details were provided for a robust assessment of this outcome. No data were provided on the distances between subjects and speakers or the direction of the speakers relative to the location of the neighbor's territory. No details were provided about stimulus donors or efforts to ensure that non-neighbors were indeed strangers in the sense of being unfamiliar to subjects. No information was provided on the relative timing of the two alternatives or how frequently each was broadcast. Finally, no quantitative data were provided regarding the numbers of advertisement calls, aggressive calls, or movements males made in response to the two stimuli. Despite these missing details, the results reported by Bourne et al. (2001) certainly suggest the presence of a dear enemy effect. Data from more recent playback experiments based on measuring aggressive thresholds, as described in Sect. 7.3.6 for *Babina adenopleura* (Chuang et al. 2017), indicate that territorial males have higher thresholds for responding aggressively toward their nearby neighbors (J. Tumulty and M. A. Bee in preparation). While still preliminary, research on golden rocket frogs is exciting because it would imply the likely independent evolution of a dear enemy effect in anurans (Fig. 7.1).

7.3.3 Agile Frogs (Ranidae)

7.3.3.1 Agile Frogs: Natural History

The agile frog, *Rana dalmatina*, occurs widely throughout Europe. The species is often described as "territorial" in the literature (e.g., Lesbarrères and Lodé 2002; Lodé and Lesbarrères 2004; Lodé et al. 2005; Lesbarrères et al. 2008; Lodé 2009); however, there appears to be some uncertainty about this species' territorial behavior and indeed its mating system. For example, in their study of the dear enemy effect, Lesbarrères and Lodé (2002) have little to say about, nor do they cite other authorities on, the territorial behavior of their study species, aside from stating that "defence behaviour is generally strict." They do not describe overt aggressive behaviors nor do they describe any distinct vocalizations, such as aggressive calls, used in territorial defense. Schneider et al. (1988) speculated that one of two call types described in their study might serve some purpose in territoriality, but this was not confirmed. In one study, the agile frog's breeding season was reported to last only about 10 days (Lodé and Lesbarrères 2004), which would place them toward the explosive end of the breeding continuum in anurans, as had already been noted by Wells (1977). Indeed, Lesbarrères et al. (2008) considered them an explosive breeder. Explosive breeders generally do not defend breeding resources. Other studies have reported longer breeding periods lasting 35–55 days (e.g., Schneider et al. 1988). How best to describe the species' mating system is unclear. For example, Lodé and Lesbarrères (2004) make direct or indirect reference to agile frogs as being monogamous, polygynous, monandrous, and polyandrous (see also Lodé et al. 2005). This apparent confusion testifies, perhaps, to the lability of the mating behaviors of some anurans. For our purposes, it is unclear whether agile frogs actually exhibit aggressive territorial behaviors similar to those described earlier for bullfrogs and golden rocket frogs. Males produce advertisement calls consisting of a series of about 25 pulse groups, each composed of 3–12 pulses (Fig. 7.3c) (Schneider et al. 1988; Lesbarrères and Lodé 2002). Some authors have described choruses in this species (Schneider et al. 1988), while others claim the species does not form choruses (Lodé and Lesbarrères 2004). Lesbarrères and Lodé (2002) report that this species' advertisement calls are individually distinctive, but this assessment has been questioned on methodological grounds (Bee et al. 2016). Uncertainty about the breeding ecology and vocal and aggressive behaviors of agile frogs serve to qualify interpretations of results from studies of the dear enemy effect in this species.

7.3.3.2 Agile Frogs: Neighbor Recognition

In their field playback experiment, Lesbarrères and Lodé (2002) presented subjects with "familiar and unfamiliar calls" and measured the acoustic properties of their response calls. The authors reported that males produced significantly longer calls, consisting of more pulse groups, in response to hearing unfamiliar calls (Fig. 7.6), and concluded this was evidence for neighbor recognition and the dear enemy effect. A number of methodological issues, however, raise significant doubts about this interpretation. The most significant issue is that the relative increase in call duration in response to unfamiliar calls compared to familiar calls was small. In response to familiar calls, call duration was about 4.1 s, ranging from about 1.4 to 13.8 s (Fig. 7.6). In response to unfamiliar calls, call duration averaged 5.2 s and ranged from 2.0 to 15.8 s. Thus, the stated evidence for the dear enemy effect is that males produced calls that were, on average, about 1 s longer in response to hearing unfamiliar calls compared with familiar calls. The authors provided no evidence to suggest that such a small increase in call duration is indicative of territorial aggression. Other frogs are known to lengthen their calls in other, nonaggressive contexts, such as increasing their attractiveness to females (Reichert and Gerhardt 2012) and matching the length of calls produced by other nearby males (Gerhardt et al. 2000). The authors make no note of overt aggressive responses in the form of distinct aggressive calls, aggressive behavioral displays, or aggressive movements toward the speaker. In short, little evidence was provided to show that males behaved less aggressively in response to hearing familiar calls compared with unfamiliar calls.

This study is plagued by other problems too. There was no indication that familiar calls were those of a male's adjacent territorial neighbor. Instead, familiar calls were "from a male of the same pond," whereas unfamiliar calls were "from a male of a different pond" (p. 289, Lesbarrères and Lodé 2002). This represents a significant methodological departure from nearly all previous studies of vocally mediated neighbor recognition. Stimuli were played back "in the immediate vicinity (range 1–3 m)" of subjects (p. 289, Lesbarrères and Lodé 2002). Unfortunately, no information was provided about typical intermale distances nor was it specified whether both types of calls were broadcast from the same location or whether the location of stimulus broadcasts was somehow related to the location of a subject's actual neighbors. This too represents a departure from methodological norms for

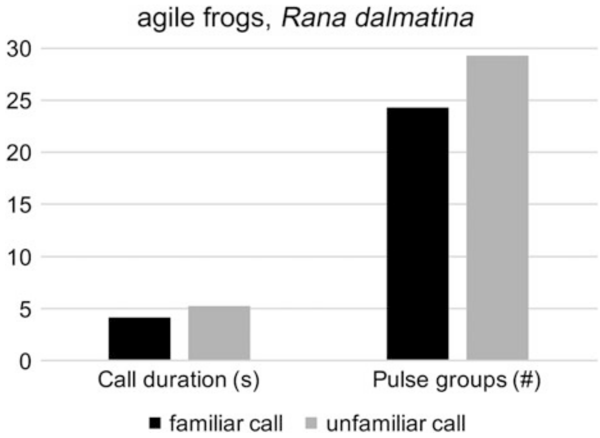


Fig. 7.6 Results from a playback study of the dear enemy effect in the agile frog (*Rana dalmatina*). Results shown are redrawn from Lesbarrères and Lodé (2002) and depict the mean duration and number of pulse groups of calls produced during playbacks of the calls of a presumed familiar male and an unfamiliar male

acoustic playback studies of neighbor recognition. Playback amplitudes varied across a 7 dB range, but no assurance was given that this range was equal for both stimulus types. Is it possible that unfamiliar stimuli were broadcast at relatively higher average sound pressure levels? One additional weakness shared with many other studies of neighbor recognition (e.g., Bee 2003b) is that no information was provided on the acoustic differences between familiar and unfamiliar calls used as stimuli. Might sampling error have resulted in acoustic differences in the stimuli that were unrelated to individual identity but nevertheless sufficient to elicit the small differences observed in behavioral responses? This is impossible to assess considering that the authors do not report how many stimulus donors were used for the familiar and unfamiliar calls.

Considering the totality of the problematic issues reviewed in this section, it remains unclear at present whether males of this species are territorial, whether they recognize their territorial neighbors by voice, and whether they exhibit relatively lower levels of aggression toward familiar neighbors. The study by Lesbarrères and Lodé (2002), specifically, illustrates the problems that can arise when experimental tests of neighbor recognition get ahead of basic natural history and when standard disciplinary practices are not followed. Additional efforts to reevaluate their results using more robust observational, experimental, and analytical methods could prove informative and worthwhile.

7.3.4 Strawberry Poison Frogs (Dendrobatidae)

7.3.4.1 Strawberry Poison Frogs: Natural History

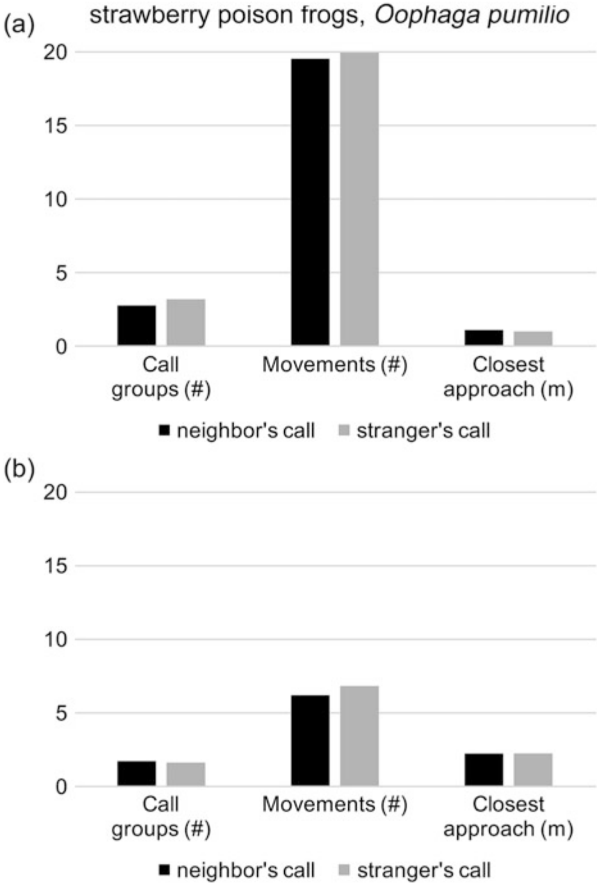
Among the diversity of brightly colored poison frogs in the family Dendrobatidae is the well-studied strawberry poison frog, *Oophaga* (formerly *Dendrobates*) *pumilio*, a species common throughout Central America. Strawberry poison frogs are small and diurnal, inhabit the leaf litter on the forest floor, and exhibit well-developed parental care. Across their geographic range, populations of these frogs vary in aposematic coloration from red, to green, to dark blue (Summers et al. 1999; Maan and Cummings 2012). The social and reproductive behaviors of strawberry poison frogs have been studied extensively (Bunnell 1973; McVey et al. 1981; Donnelly 1989a, b, c; Pröhl 1997a, b, 2002, 2003, 2005a; Pröhl and Hödl 1999; Summers et al. 1999; Pröhl and Berke 2001; Haase and Pröhl 2002; Maan and Cummings 2012; Meuche et al. 2012; Crothers and Cummings 2015). Early work by Donnelly (1989a, b) and Pröhl (1997b) reported that males establish and aggressively defend multipurpose territories. These territories were reported to include foraging sites, shelter, courtship areas, oviposition sites, and tadpole-rearing sites, although more recent work has questioned the defense of breeding resources and suggests males may simply defend areas with good sound transmission characteristics and where they are likely to encounter females (Pröhl 2005b). Territory defense can include vocalizations, phonotaxis toward intruders, brief chases, visual push-up displays, and, in the most escalated contests, physical fights. While early accounts tended to characterize the mating system of this species as resource defense polygyny, more recent treatments suggest a lek mating system may better describe its reproductive behavior (Pröhl 2005b; Wells 2007). Whatever it is that males defend, individual males often defend the same area for periods lasting a few weeks or months to perhaps as long as a few years (Pröhl 1997b; Pröhl and Hödl 1999; Pröhl and Berke 2001). Thus, field observations indicate that males aggressively defend long-term territories and potentially share boundaries with nearby neighbors for extended periods of time, suggesting the hypothesis that vocally mediated dear enemy behavior could be adaptive. Pröhl (2003) demonstrated that advertisement calls (Fig. 7.3d) in strawberry poison frogs are individually distinct.

7.3.4.2 Strawberry Poison Frogs: Neighbor Recognition

Bee (2003b) conducted two field playback experiments in a population of strawberry poison frogs on Isla Bastimentos, in Panama, to test the hypothesis that territorial males behaviorally discriminate between the calls of strangers and a nearby neighbor. Both experiments consisted of broadcasting the prerecorded calls of either a neighbor (from an adjacent territory) or a stranger (from a different part of the island). In experiment 1 ($N = 24$), the speaker was located approximately halfway between the position of the subject and the estimated center of the neighbor's territory. Experiment 2 ($N = 22$) was similar save for one main difference: the playback speaker was positioned at the approximate center of the neighbor's territory. Four response variables were measured to quantify a subject's aggressive behavior. These included the number of call groups produced, the number of movements made, the subject's closest approach to the speaker, and the maximum absolute distance the subject approached toward the speaker. Although previous studies had reported the use of aggressive calls (Bunnell 1973; Zimmermann 1990) and visual displays (Baugh and Forester 1994) during aggressive encounters in strawberry poison frogs, these measures were not quantified by Bee (2003b). Push-up displays were not observed, and the differences between advertisement and aggressive calls had not been well quantified or described at the time (but see Meuche et al. 2012) nor were they acoustically apparent during the experiment.

The results of both experiments were unequivocal in their lack of support for a dear enemy effect. In experiment 1, when the speaker was located midway between the centers of the subject's and neighbor's territories, subjects exhibited clear aggressive responses to playbacks. They oriented to the speaker and then approached it while calling at an elevated rate, and often continued hopping around the vicinity of the speaker, as if searching for an intruder (Fig. 7.7a). The behavior observed during playback was similar to that observed during natural aggressive interactions between two individuals. However, there was no evidence that subjects responded differently to playbacks of the calls of neighbors and strangers (Fig. 7.7a). One plausible explanation for this outcome was that subjects perceived

Fig. 7.7 Results from a playback study of the dear enemy effect in the strawberry poison frog (*Oophaga pumilio*). Results shown are redrawn from Bee (2003b) and depict the mean numbers of call groups and movements and the closest distance approached toward the playback speaker during broadcasts of the calls of a neighbor or a stranger from (**a**) midway between the subject's and neighbor's territory or from (**b**) the center of the neighbor's territory



the calls of both neighbors and strangers as being equally threatening (and hence deserving of equal responses). That is, subjects may have responded equally aggressively to the neighbor and stranger stimuli because both were perceived as being too close to the subject's calling position. To evaluate this possibility, experiment 2 repeated experiment 1 with the speaker located in the approximate center of the neighbor's territory. In this second experiment, males also responded to playbacks, though their responses were generally much weaker than those observed in experiment 1. And again, there was no evidence to suggest that males reacted differently in response to hearing the calls of neighbors versus strangers (Fig. 7.7b).

Together, the results of both experiments suggested that male strawberry poison frogs readily responded to the calls of neighbors and strangers, but did not behaviorally discriminate between them. Why might this be the case, given that males defend long-term territories and their calls are individually distinctive? At an ultimate level, Bee (2003b) speculated that both neighbors and strangers might represent equal threats to a territory holder. According to Temeles (1994), this might well be the case if territories do not, in fact, include limited breeding resources, such as oviposition sites and tadpole-rearing sites (Pröhl 2005b). Neighbors might be especially threatening if they can disrupt the prolonged courtship involved in this species' mating behavior. In addition, the costs of repeated aggression might be sufficiently low in strawberry poison frogs so as to offset any benefit of an evolved recognition system for dear enemy behavior. In an extensive field study by Pröhl and Berke (2001), most aggressive interactions were settled with vocal displays at shared territory borders. Aggressive encounters rarely escalated to fighting (just 17 % of interactions), and territory takeovers were even more rare (just 3 % of interactions). Potential costs of aggression related to increased exposure to predators, or decreased vigilance against predators, might be relaxed because of these animals' aposematic coloration. Moreover, some studies have reported prior residence effects, in that aggressive encounters were typically won by the territory holder (Baugh and Forester 1994; Pröhl and Berke 2001). Thus, instead of a dear enemy effect, evolution may have struck upon a low-cost "residents always win" solution to reduce the costs of repeated aggressive interactions between neighbors.

At a more proximate level, Bee (2003b) speculated that male strawberry poison frogs might be unable to perceptually discriminate the magnitudes of differences in acoustic call properties that *reliably* distinguish among individuals. The key to this argument concerns the reliability of individual variation under natural conditions. Several individually distinctive call properties examined by Pröhl (2003)—including the two properties with the highest CV_a/CV_w ratios and repeatability estimates (pulse rate and call duration)—varied significantly with temperature. Therefore, individual differences in vocalizations might often be obscured by microhabitat variation in temperature. As a result, the same individual might sound different, for example, when calling from different perches in his territory exposed to different temperature regimes. Thus, the temperature dependency of some call properties might render them less reliable vocal cues for discriminating among individuals. One acoustic property that was not temperature dependent, but was individually distinctive, was the dominant peak of the call's frequency spectrum. Given its frequency between 4 and 5 kHz (Fig. 7.3d), the dominant frequency of the call most likely falls within the range of the basilar papilla of the anuran inner ear. This structure is tuned to high frequencies and lacks the tonotopic organization characteristic of the anuran amphibian papilla. At such high frequencies, the basilar papilla probably cannot resolve differences in frequency independently of differences in sound level, which would be expected to vary with distance from the source (Gerhardt and Huber 2002). Therefore, it seems likely that the strawberry poison frog's auditory system might constrain the ability to reliably discriminate among males based on individual differences in dominant frequency. However, as was the case for golden rocket frogs (see Sect. 7.3.2), and as we will see in the next section, high-frequency calls do not necessarily exclude the possibility of perceptual discrimination between individuals.

7.3.5 Concave-Eared Torrent Frogs (Ranidae)

7.3.5.1 Concave-Eared Torrent Frogs: Natural History

Among anurans, the concave-eared torrent frog, *Odorrana tormota* (formerly *Amolops tormotus*), possesses two extremely rare vocal characteristics that make it an interesting species to study. It has an exceptionally complex vocal repertoire, including calls characterized by the presence of several nonlinear acoustic phenomena (e.g., frequency jumps, subharmonics, biphonation, and deterministic chaos), and it uses ultrasonic frequencies in communication (Feng et al. 2002, 2006, 2009b; Suthers et al. 2006). A male's vocal repertoire consists of eight different call types (Feng et al. 2002, 2009b; Narins et al. 2004). The behavioral functions of these different call types remain to be investigated. Feng et al. (2009b) reported individual vocal distinctiveness in the species' "long calls" (Fig. 7.3e), which were only produced during robust choruses. The species breeds along noisy streams in the mountains and hills of two small regions of eastern China. It is believed that ultrasonic frequencies allow these frogs to communicate over the din of low-frequency noise generated by fast flowing water. Although the vocal behavior of this frog has been investigated extensively in a number of recent studies (Narins et al. 2004; Suthers et al. 2006; Feng and Narins 2008; Gridi-Papp et al. 2008; Shen et al. 2008, 2011; Arch et al. 2011), we still lack important details about much of the basic natural history and behavior of this species. For example, we presently know little about the species' mating system, its reproductive biology, and the extent to which males exhibit aggressive territorial behaviors similar to those described previously for bullfrogs (Sect. 7.3.1) and strawberry poison frogs (Sect. 7.3.4). In this respect, studies of the dear enemy effect in concave-eared torrent frogs are built upon the same uncertain foundation that underlies previous studies of the phenomenon in agile frogs (Sect. 7.3.3).

7.3.5.2 Concave-Eared Torrent Frogs: Neighbor Recognition

Feng et al. (2009a) investigated the dear enemy effect in concave-eared torrent frogs. In a laboratory playback experiment, they presented male subjects with the calls of neighbors and strangers and compared the number of response vocalizations evoked by these two stimuli relative to each other and to a control condition in which no sounds were broadcast. During the silent control condition, males did not call. As shown in Fig. 7.8, subjects gave more evoked vocal responses upon hearing the call of a stranger compared with hearing the call of a neighbor. Moreover, more of their calls were produced antiphonally with the stimulus in response to the calls of strangers compared with neighbors (Fig. 7.8). Feng et al. (2009a) also report that males were more agitated during playback of strangers' calls, displaying various motor activities that were interpreted as aggressive responses. Interestingly, only 64 % of subjects responded with relatively more evoked calls in response to

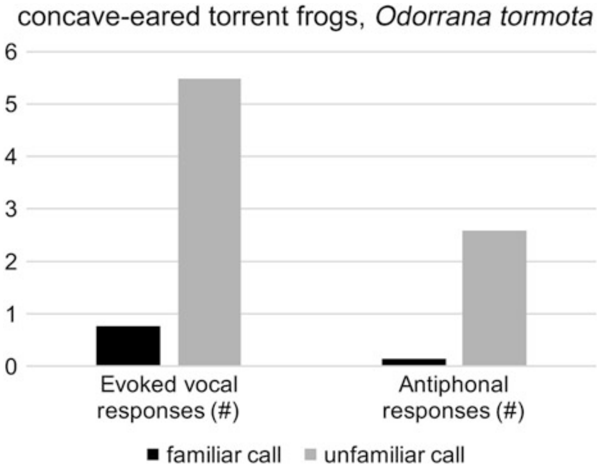


Fig. 7.8 Results from a playback study of the dear enemy effect in the concave-eared torrent frog (*Odorrana tormota*). Results shown are from Feng et al. (2009a) and depict the mean number of evoked vocal responses and antiphonal calls produced during playback of the calls of a presumed familiar male and an unfamiliar male

strangers compared with neighbors. The other 36 % of subjects treated the two calls as being equivalent. As pointed out by Feng et al. (2009a), this is noteworthy because it is close to the correct classification success of 54.6 % resulting from a DFA of male calls (Feng et al. 2009b). Thus both real males and the DFA returned similar rates of classification errors, perhaps due in both cases to a somewhat low level of individual distinctiveness in male calls compared with that reported for other species.

On the surface, these results with concave-eared torrent frogs appear consistent with neighbor recognition and the dear enemy effect. However, several aspects of methodology make this interpretation somewhat less certain. In stark contrast to previous studies of neighbor recognition in bullfrogs, golden rocket frogs, agile frogs, and strawberry poison frogs, all of which tested subjects under natural conditions in the field, Feng et al. (2009a) conducted their playback experiment in the laboratory. Subjects were tested in small, plastic tanks, and evoked vocalizations were quantified in response to stimuli broadcast from directly above the test tank. This methodology eliminates the defended resource, potentially confounding the biological distinction between neighbors and strangers. The evoked vocal response of this species has not been directly linked to overt aggressive behaviors, as opposed to increased competition to attract females (see Bee et al. 2016 for a more in-depth discussion of this general problem). In addition, to my knowledge, the species has not been shown to defend long-term territories that contain some kind of breeding resource (as opposed to more short-term calling sites). As in the study of agile frogs (Lesbarrères and Lodé 2002), the study by Feng et al. (2009a) also used an atypical criterion to define which males were neighbors versus strangers. A neighbor was considered to be a male observed calling within the same general area as the subject (<20 m × 20 m). Males that were calling greater than 60 m apart were considered strangers. Feng et al. (2009a) argued that males called at amplitudes that were sufficiently high to render calls audible to other males in the same small area, but not to males separated by larger distances. Thus, in their study, an audible frog was apparently considered a neighboring frog. Most previous studies, in contrast, have used the male occupying an immediately adjacent territory as the donor of a neighbor stimulus. Finally, no acoustic analyses of the neighbor and stranger stimuli were provided to rule out the possibility that these two classes of stimuli differed along some dimension other than individual identity, such as the extent to which they conveyed aggressiveness or were sexually attractive to females. Additional work on concave-eared torrent frogs is needed before more firm conclusions can be drawn about whether this species exhibits neighbor recognition and the dear enemy effect. That said, current evidence is certainly suggestive of a potential dear enemy effect.

7.3.6 Olive Frogs (Ranidae)

7.3.6.1 Olive Frogs: Natural History

The olive frog, *Babina* (formerly *Rana*) *adenopleura*, is common throughout China and Taiwan (Chuaynkern et al. 2010). A recent field study by Chuang et al. (2013) found that males of this species defend territories that function as oviposition sites used by females. There is considerable variation in male mating success in this species (M-F Chuang, MA Bee, and Y-C Kam unpublished data). Hence, as for bullfrogs, the mating system is best described as resource defense polygyny. Individual males have been observed to defend the same territory for up to 43 consecutive nights (Chuang et al. 2013). Territory residents thus potentially share boundaries with neighbors over periods of several days and perhaps up to several weeks, although this has yet to be quantified. The species has a vocal repertoire of six distinct call types, two of which (the “territorial call” and the “encounter call”) are used in aggressive encounters between males (Chuang et al. 2016). Thus, in terms of its reproductive and vocal behavior, male olive frogs are similar in many ways to male bullfrogs (Sect. 7.3.1). The olive frog advertisement call consists of a series of four or five notes produced in rapid succession in just under 1 s (Fig. 7.3f). Chuang et al. (2017) have reported that advertisement calls are individually distinctive.

7.3.6.2 Olive Frogs: Neighbor Recognition

Chuang et al. (2017) conducted a field playback experiment to test the hypothesis that territory residents are less aggressive in response to hearing the call of a territorial neighbor compared with hearing a stranger. They used a novel methodological approach to test this hypothesis. The approach was based on determining a territory resident’s “aggressive threshold” in response to neighbors and strangers. Previous field playback studies of wrinkled toadlets (*Uperoleia rugosa*; Myobatrachidae), Pacific treefrogs (*Pseudacris regilla*; Hylidae), and spring peepers (*Pseudacris crucifer*; Hylidae) measured aggressive thresholds as the minimum sound pressure level required to elicit an aggressive response from males defending calling sites (Robertson 1984; Rose and Brenowitz 1991; Brenowitz and Rose 1994; Marshall et al. 2003). The logic behind using this approach to investigate neighbor recognition in olive frogs was that, in the context of the dear enemy effect, territory holders should have relatively higher aggressive thresholds in response to neighbors compared with strangers. Stated differently, territory holders should be more tolerant of neighbors and less tolerant of strangers.

During playbacks, the acoustic stimulus started at a low level and was repeated on a 1 min loop. On each successive broadcast, its amplitude was increased by 2 dB. This incremental process continued until the subject responded with territorial calls, encounter calls, or both (or until the technical limits of the playback system were reached). Ten males heard the previously recorded call of their nearest territorial neighbor as the stimulus. Another group of ten males heard the call of a stranger, with each subject in this group tested with a different stranger stimulus. There were no differences between the acoustic properties of the neighbor and stranger stimuli. There were, however, significant differences in aggressive thresholds in response to these two stimuli. The mean aggressive threshold for eliciting territorial calls was significantly higher, by about 4 dB, in response to the calls of neighbors compared with strangers (Fig. 7.9). The most aggressive male in the group presented with their neighbor's calls had a threshold that was about 7 dB higher than the most aggressive male in the group that heard the calls of strangers (Fig. 7.9). Thresholds for eliciting encounter calls were higher than those for territorial calls, suggesting it is an aggressive vocalization used in more escalated encounters. There were no differences in the aggressive thresholds for eliciting encounter calls in response to neighbors and strangers; however, the lowest threshold in each experimental group was about 5 dB higher in the group that heard their neighbors. The received sound pressure level of another individual's calls should generally correlate inversely with

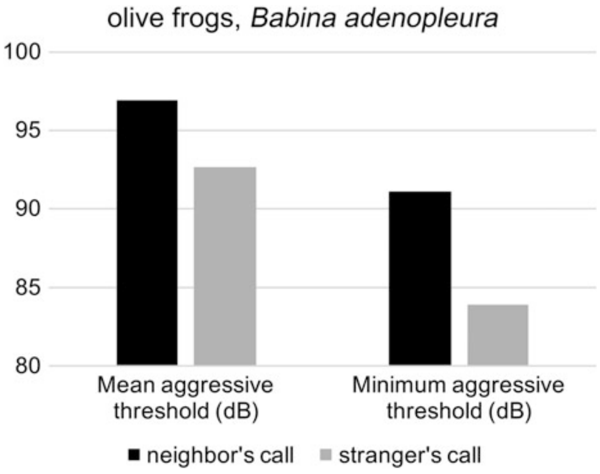


Fig. 7.9 Results from a playback study of the dear enemy effect in the olive frog (*Babina adenopleura*). Results shown are redrawn from Chuang et al. (2017) and depict the mean and minimum aggressive thresholds, defined as the sound pressure level required to elicit a territorial call, in response to playbacks of the calls of a neighbor and a stranger

distance to the source due to spherical spreading and other influences on sound propagation (Wiley and Richards 1978; Richards and Wiley 1980; Boatright-Horowitz et al. 1999). Thus, lower aggressive thresholds in response to a stranger's calls should translate into a lower tolerance for encroachment by a stranger compared with a neighbor. Taken together, the results of Chuang et al. (2013, 2016, in press) indicate that male olive frogs are territorial, defend breeding resources, have individually distinct advertisement calls, and are less aggressive in response to hearing the calls of their nearby neighbors compared with those of strangers.

7.3.7 Section Summary

For those interested in comparative social cognition, anurans present some exciting opportunities to investigate the mechanisms and evolution of social recognition. This is because some anuran species (e.g., prolonged breeders with resource defense mating systems) appear to exhibit social recognition of neighbors, while others (e.g., many explosive breeders and species with lek polygyny mating systems) almost certainly do not. Current evidence indicates that bullfrogs (Ranidae), golden rocket frogs (Dendrobatidae), and olive frogs (Ranidae) exhibit social recognition in the form of a dear enemy effect, while others, such as the strawberry poison frog (Dendrobatidae), do not. For some other species, such as agile frogs (Ranidae) and concave-eared torrent frogs (Ranidae), the data are perhaps suggestive of a possible dear enemy effect, but they are also less conclusive due to various methodological issues, and other interpretations remain plausible. One study not reviewed in this section suggests the presence of a dear enemy effect in another territorial species, the green frog (*Rana clamitans*, Ranidae) (Owen and Perrill 1998), which is a close relative of bullfrogs and has similar breeding behavior; however, that study did not distinguish between the phenomenon itself and one hypothesized mechanism for the behavior, making interpretation of its results somewhat difficult (Bee and Schachtman 2000).

Unfortunately, broader conclusions about the prevalence of neighbor recognition in anurans and the ecological and social factors that favor its evolution are not possible at this time. The current data suggest factors related to the defense of breeding resources, such as oviposition sites, may be important, as originally suggested by Temeles (1994). From a comparative perspective, the presence of neighbor recognition in the Ranidae and Dendrobatidae is suggestive of independent origins, as these two families diverged about 160 million years ago (Fig. 1; Hedges and Kumar 2009). Even within the Ranidae, there have likely been multiple origins of neighbor recognition, as bullfrogs and olive frogs last shared a common ancestor some 41 million years ago. Given the evolutionary lability of frog social and reproductive behaviors, multiple evolutionary origins of territorial neighbor recognition would not be surprising. Additional comparative studies using standardized methods (e.g., measuring aggressive thresholds) across a wider range of species are needed to test these ideas. One exciting implication of multiple evolutionary origins would be the possibility of diversity in underlying perceptual and cognitive mechanisms. These mechanisms have so far been investigated in only one species, however. It is these studies that are reviewed in the next section.

7.4 Mechanisms of Neighbor Recognition in Bullfrogs

In this section, I review studies that have elucidated the acoustic, perceptual, and cognitive bases of neighbor recognition in North American bullfrogs. Recall that Davis (1987) showed that territorial male bullfrogs perceptually and behaviorally discriminate between the advertisement calls of neighbors and strangers, responding more aggressively to the calls of strangers (Fig. 7.5). Territorial males were also shown to associate the call of a familiar neighbor with the correct location of that neighbor's territory. In many ways, Davis' (1987) results indicated some level of evolutionary convergence in the adaptations of neighbor recognition and the dear enemy phenomenon between territorial songbirds and territorial frogs. Efforts to dig deeper into understanding the nature of this adaptation in bullfrogs have taken several important next steps: (1) describe how patterns of individual variation in the acoustic properties of advertisement calls give rise to individual vocal distinctiveness (Sect. 7.4.1), (2) empirically demonstrate which acoustic properties receivers use in perceptually discriminating the voices of neighbors from strangers (Sect. 7.4.2), and (3) identify cognitive processes that enable territorial bullfrogs to learn about, and exhibit reduced aggression toward, their neighbor's calls (Sect. 7.4.3).

7.4.1 Acoustic Basis of Neighbor Recognition in Bullfrogs

Each note in a bullfrog advertisement call (Figs. 7.3a and 7.10) has a bimodal frequency spectrum consisting of a series of harmonics. The lower-frequency mode is centered between 200 and 400 Hz, with the dominant frequency of the call near 200 Hz; the higher-frequency mode, which has a somewhat lower relative peak amplitude, contains sound energy between 1000 and 2000 Hz (Fig. 7.10a). These two modes match the tuning of the two sensory papillae in the bullfrog's inner ear, the amphibian papilla and basilar papilla, respectively (reviewed in Gerhardt and Huber 2002; Narins et al. 2007; Simmons 2013). The fundamental frequency itself varies between about 90 and 130 Hz but contains relatively little acoustic energy, making it a so-called "missing fundamental," which auditory nerve fibers nevertheless encode in a robust pattern of precisely timed action potentials (Simmons and Ferragamo 1993). The fundamental frequency is clearly evident in the call's quasiperiodic fine temporal structure (see inset in Fig. 7.10b) and varies strongly, and inversely, with body size ($R^2 = 0.81$) (Bee and Gerhardt 2001b).

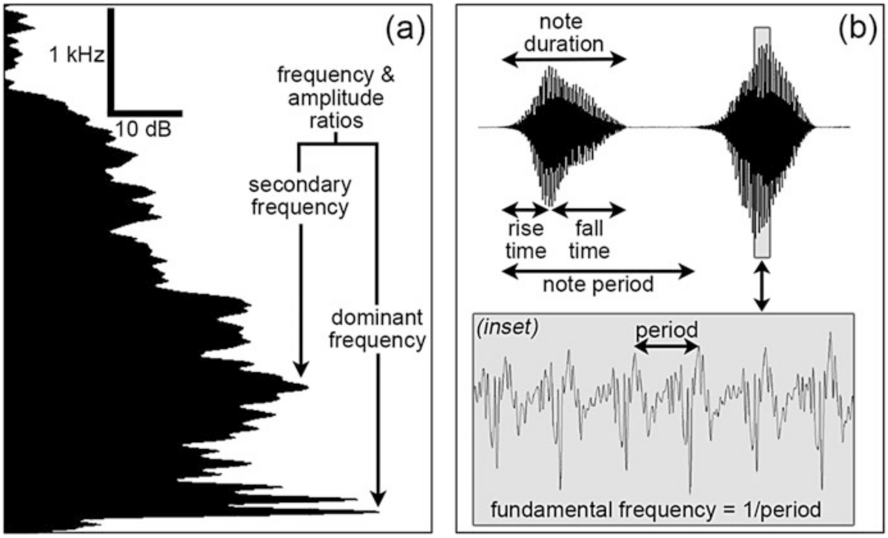


Fig. 7.10 Acoustic properties of bullfrog (*Rana catesbeiana*) advertisement calls. Shown here are the (**a**) spectral and (**b**) temporal properties of bullfrogs calls (illustrated here as a two-note call) measured by Bee and Gerhardt (2001b) and Bee (2004). Duty cycle was additionally calculated as note duration divided by note period (i.e., the proportion of the note period taken up by the duration of a note). Measures of individual variation in these properties are reported in Table 7.1

A published abstract and unpublished dissertation by Haas (1976, 1977) reported that several acoustic properties of bullfrog advertisement calls were individually distinct, but sample sizes were small and minimal statistical analyses were performed. This finding was later confirmed and extended by Bee and Gerhardt (2001b), who investigated individual vocal distinctiveness in a sample of nearly 1100 advertisement calls recorded from 27 males, with calls recorded from each individual on two different occasions separated by 1–16 days. These analyses focused on ten spectral and temporal properties (Table 7.1; Fig. 7.10), including fundamental frequency, dominant and secondary frequencies (defined as the harmonics of greatest relative amplitude in the lower and upper frequency modes of the call), and measures of their frequency and amplitude ratios, note duration, the onset and offset ramps of a note (i.e., rise time and fall time, respectively), note period, and note duty cycle (note duration/note period). Regression analyses were used to remove variation in call properties due to water temperature while retaining important sources of individual variation related to body size.

All of the acoustic properties measured by Bee and Gerhardt (2001b), and illustrated in Fig. 7.10, varied significantly more among individuals than within individuals. All ratios of among-individual to within-individual coefficients of variation (CV_a/CV_w) were greater than 1.0 (Table 7.1). Individual variability was also temporally stable, as all call properties were also “repeatable” across recordings made on different days, with intraclass correlation coefficients ranging between 0.61 and 0.93 (Table 7.1). The fundamental frequency (and the correlated

Table 7.1 Measures of coefficients of variation (CV as %) for the acoustic properties of bullfrog advertisement calls (see Fig. 7.10) determined among individuals (CV_a) and within individuals (CV_w, averaged across individuals) as well as their ratio and repeatability (intraclass correlation coefficient) measured by comparing two different recording sessions separated by 1–16 days ($N = 27$); also shown are estimates of just-meaningful differences (JMD as a %) for several properties

Call property	CV_a	Mean CV_w	CV_a/CV_w	Repeatability	JMD
Fundamental frequency	7.0	1.5	4.7	0.93	4.0
Dominant frequency	7.0	1.5	4.7	0.93	5.0
Secondary frequency	10.5	5.6	1.9	0.51	6.0
Frequency ratio	11.1	5.8	1.9	0.44	6.0
Relative amplitude	27.0	19.1	1.4	0.71	25.0
Note duration	8.9	6.6	1.4	0.56	10.0
Rise time	18.0	12.6	1.4	0.63	21.0
Fall time	13.8	12.4	1.1	0.61	17.0
Note period	9.1	5.6	1.6	0.43	–
Duty cycle	8.0	6.2	1.3	0.61	–

Data on CVs and repeatabilities from Bee and Gerhardt (2001b); data on JMDs from Bee (2004)

harmonic frequencies, especially dominant frequency) was the most individually distinctive properties. For example, fundamental frequency had a CV_a/CV_w ratio of 4.7 and an intraclass correlation coefficient of 0.93. The high CV_a/CV_w ratio stemmed in large part from the fact that fundamental frequency exhibited very little variation within individuals, having a mean CV_w of just 1.5 % and a range of 0.6–3.9 %. (Given a mean fundamental frequency of 109 Hz, a mean CV_w of 1.5 % corresponds to typical within-individual variation less than ±2 Hz.) Several different DFAs were performed, all of which indicated calls were assigned to the correct individual significantly more often than expected by chance. In the most biologically realistic of these DFAs, 1000 random samples of five males each were drawn from the sample of 27 males, and correct classification was computed based on using calls recorded on two different nights as the training and test datasets. The biological realism of these analyses stems from the fact that territorial male bullfrogs never have as many as 27 neighbors at a time (more often they have just two or three), and individually distinctive acoustic properties should be stable over time. In this analysis, 67 % of calls were correctly assigned to the individual that produced them. The first canonical root of these DFAs typically explained 70–80% of the variation in calls and was strongly correlated with fundamental frequency. Bee (2004) later showed that these general patterns of individual variation were maintained when the variability across the multiple notes within individuals' calls (Fig. 7.3a) was also taken into account. That study showed in addition that individually distinct patterns of within-call variation could potentially be used to distinguish among some individuals. The clear take-home message from these acoustic and statistical analyses was that bullfrog advertisement calls are individually distinct. More important, these analyses provided a strong foundation for making predictions about the perceptual basis of neighbor recognition.

7.4.2 Perceptual Basis of Neighbor Recognition in Bullfrogs

We might generally expect the perceptual basis of vocally mediated social recognition to involve discrimination of individual differences occurring in a select subset of signal properties. For example, signal properties differ in the extent to which among-individual variation in a population exceeds typical magnitudes of within-individual variation. Properties with the highest CV_a/CV_w ratios should provide receivers with the best information about a signaler's individual identity and, therefore, might be particularly useful for discriminating among individuals. Even with statistically significant among-individual variation, however, small absolute differences among individuals in some properties may have no biological function in recognition if the receiver's perceptual system cannot resolve them, that is, if individual differences are smaller than a receiver's JND (Chap. 6). Even if individual differences exceed a receiver's JND, they may still fail to evoke a discriminative behavioral response because they are smaller than the animal's JMD. In addition, not all signal properties propagate equally well through the environment (Wiley and Richards 1978; Richards and Wiley 1980); therefore, we might expect receivers to discriminate among individuals using signal properties that propagate with the least attenuation and degradation. Along these same lines, we should generally expect receivers to heavily weight signal properties for which neural encoding is resistant to the effects of noise.

Based on these and other considerations, Bee and Gerhardt (2001b) hypothesized that territorial male bullfrogs should use individual differences in fundamental frequency (or correlated spectral properties) to perceptually and behaviorally discriminate between neighbors and strangers. Fundamental frequency, and the correlated property dominant frequency, exhibited the smallest variation within individuals and the greatest relative variation among individuals (Bee and Gerhardt 2001b). Both of these acoustic properties propagate well in the species' natural breeding habitat across distances that typically separate adjacent territorial males (Boatright-Horowitz et al. 1999). The bullfrog peripheral auditory system encodes fundamental frequency and the low-frequency spectral components of harmonically complex signals (e.g., dominant frequency) under a wide range of signal-to-noise conditions (Schwartz and Simmons 1990; Simmons et al. 1992, 1993; Simmons and Ferragamo 1993). Research with other anurans had previously shown that discrimination of individual differences in fundamental frequency and spectral call properties influences female mating decisions (reviewed in Gerhardt and Schwartz 2001) and the decisions males make about escalating aggressive contests (reviewed in Bee et al. 2016). So frequency is often a behaviorally salient acoustic feature of calls in anurans.

Bee and Gerhardt (2001a, c, 2002) performed a series of field playback experiments with territorial males to test two key predictions of the hypothesis that discrimination between neighbors and strangers is mediated by individual differences in fundamental frequency (or correlated spectral properties). First, using synthetic acoustic stimuli that effectively mimicked bullfrog calls, they tested the prediction that territorial males would behaviorally discriminate between a familiar and an unfamiliar call differing only in fundamental frequency (and other correlated spectral properties). Second, they tested the prediction that territory residents would associate a familiar fundamental frequency with a particular location. Their empirical approach was based on using habituation-discrimination experiments to simulate the arrival of a new and persistently calling territorial neighbor in a formerly unoccupied position. During an initial habituation phase, a synthetic call was presented repeatedly from an unoccupied "territory" adjacent to the subject's territory. Subjects responded very aggressively to the new arrival, producing advertisement calls and aggressive calls and approaching the speaker using very noisy splash displays and, in some instances, even mounting the Styrofoam speaker platform and coming "face-to-face" with the speaker in their search for a calling intruder. Eventually, and usually over the course of several hours (and in one case, 10 hours!), males habituated to the stimulus, returned to their original calling sites, and resumed giving exclusively advertisement calls. The interpretation of this change in a male's aggressive responsiveness was that the subject had become familiar with the acoustic stimulus as a result of repeated exposure and no longer perceived it as a threat to territory ownership. (Section 7.4.3 examines habituation of aggression in more detail.) Following this habituation phase, the fundamental frequency of the stimulus, its location, or both were manipulated to create a novel stimulus that was presented during a subsequent discrimination phase. Subjects in control groups continued to hear the same stimulus in the discrimination phase to which they had been exposed during the preceding habituation phase.

Over the course of several experiments, Bee and Gerhardt (2001a, c, 2002) attempted to quantify receivers' decision rules governing behavioral discrimination between calls differing in fundamental frequency. The general prediction was that habituated males would resume responding aggressively when the frequency difference between the habituating and novel stimuli exceeded the magnitude of frequency differences typically observed to occur between the separate renditions of a call performed by the same individual bullfrog (i.e., within-individual differences). These expectations are similar to the SDT scenario outlined in Fig. 7.2b. Bee (2004) evaluated this scenario using data from actual within-individual and among-individual differences in calls. As illustrated in Fig. 7.11a, within-individual differences in fundamental frequency were generally less than about 5 %. Recall that values of CV_w across individuals were uniformly less than 4 % and averaged just 1.5 % (Table 7.1). In contrast, the magnitudes of among-individual differences in fundamental frequency were typically on the order of 5 % or greater (Fig. 7.11a). Based on these patterns of observed individual differences, the predicted decision rule, or JMD (Table 7.1), for behavioral discrimination by habituated males was that they should treat differences smaller than about 5 % as the same individual (i.e., ignore the novel stimulus as if it were a familiar neighbor), but treat differences larger than about 5 % as a different individual (i.e., attack the novel stimulus as if it were a stranger).

The resulting data suggested that, indeed, territorial males employed the predicted decision rule. Subjects in control groups, which heard additional

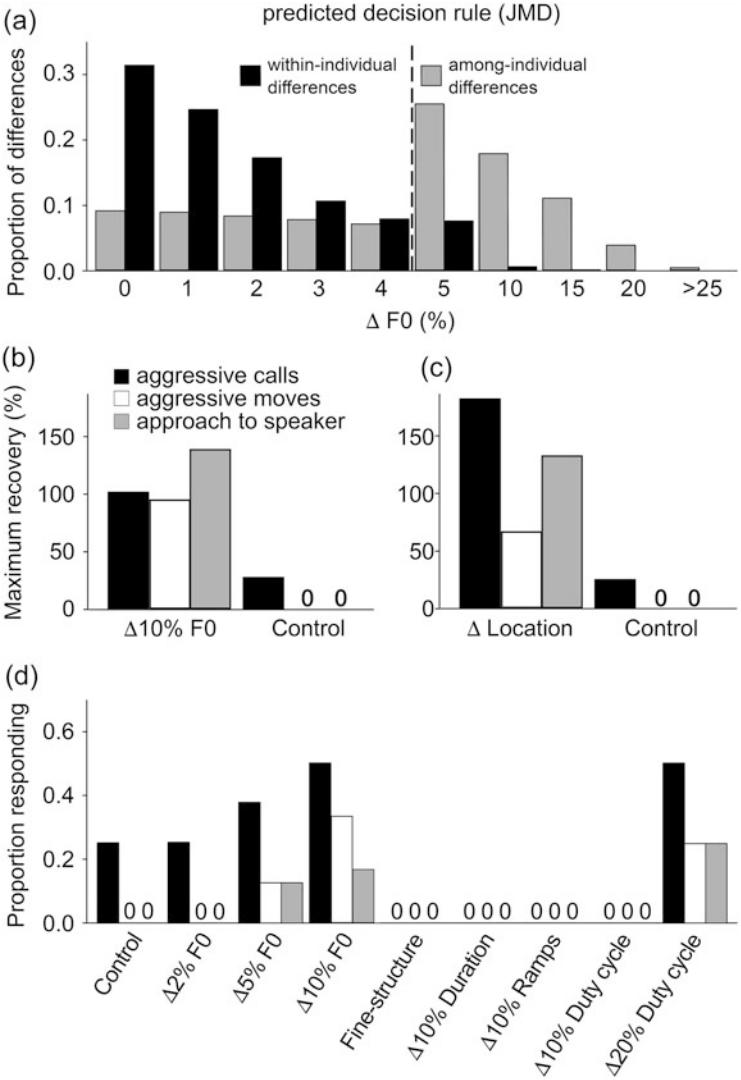


Fig. 7.11 Perceptual basis of neighbor recognition in bullfrogs (*Rana catesbeiana*). (**a**) Predicted decision rule, or just-meaningful difference (JMD), based on patterns of individual variation in fundamental frequency (F0). The histograms depict the proportions of within-individual differences and among-individual differences having the magnitudes indicated along the *x*-axis based on analyses of a large sample of recorded advertisement calls (Bee 2004). The location of the receiver's decision rule (*dashed line*) was predicted to occur near the breakpoint above which differences were more likely to reflect among-individual differences than within-individual differences. The actual data shown here reflect the signal detection theory scenario depicted in Fig. 7.2b. (**b**, **c**) Maximum recovery of habituated aggressive responses during the discrimination phase of a habituation-discrimination experiment in response to (**b**) a 10 % change in fundamental frequency (F0) and (**c**) a change in the location of the habituating stimulus (data redrawn from Bee and Gerhardt 2001c). Both 10 % changes in fundamental frequency and changes in location resulted in significant recovery of responses. (**d**) Proportions of subjects responding to changes in various acoustic properties of the novel stimulus during the discrimination phase of

presentations of the habituating stimulus during the discrimination phase, invariably exhibited little in the way of aggressive responses (Fig. 7.11b–d). If the change in fundamental frequency was 2 %, which was typical of within-individual variation, responses were similar to those of subjects in the control group (Fig. 7.11d). When males heard a novel stimulus differing by 5–10 % or more in fundamental frequency, however, they resumed giving aggressive calls and aggressively approached the playback speaker (Fig. 7.11b, d). Thus, consistent with expectations about receiver decision rules based on patterns of individual variation (Fig. 7.11a), subjects tended to ignore differences in fundamental frequency typical of within-individual variation (<5 %), but responded aggressively upon hearing a difference in fundamental frequency larger than the variation typically exhibited within individuals (≥ 5 %). Interestingly, changes imposed on the spectral envelope of the novel stimulus ("fine structure" in Fig. 7.11d), created by changing the presence/absence and relative amplitudes of specific harmonics while keeping the fundamental frequency constant, failed to elicit recovery of aggressive responses (Bee and Gerhardt 2001a). The interpretation of these results, taken together, is that territory holders learn about the fundamental frequency of a new neighbor's calls as a result of repeatedly hearing those calls across a shared boundary. In turn, learning about fundamental frequency enables them to use individual differences in this highly individually distinctive voice property as a cue for perceptually and behaviorally discriminating between at least some neighbors and strangers. Given the tight relationship between body size and fundamental frequency (Sect. 7.4.1), such a mechanism would provide a straightforward way of discriminating among neighbors and strangers based on even small differences in their size.

What about other individually distinctive voice properties? Might they also play a role in neighbor recognition? This question is difficult to answer definitively for bullfrogs based on the evidence accumulated so far. In additional habituation-discrimination experiments, 10 % changes during the discrimination phase in several temporal properties of call notes, including their duration, onset and offset ramps (i.e., rise and fall times), and duty cycle within a call, failed to elicit responses that were different from those observed in control conditions (Fig. 7.11d). Importantly, however, the magnitude of change imposed on these temporal properties (10 %) still fell within the range of CV_w values reported for those properties (Bee and Gerhardt 2001b; Bee 2004). Thus, a 10 % change in a gross temporal property (e.g., note duration) and in fundamental frequency, while equivalent in terms of percentages, was not equivalent in terms of the corresponding patterns of among-individual and within-individual differences in those properties. In other words,

◄

Fig. 7.11 (continued) habituation-discrimination experiments, including fundamental frequency (F0), harmonic fine structure, note duration, the duration of note onsets and offsets (ramps), and note duty cycle. Subjects in control groups heard additional presentations of the habituating stimulus during the discrimination phase. Only 5–10 % changes in F0 and a 20 % change in duty cycle elicited responses from more subjects than continuing playbacks of the habituating stimulus during control tests (data redrawn from Bee and Gerhardt 2001a) changes on the order of 10 % exceeded the predicted JMD for fundamental frequency, but they did not exceed the predicted JMD for gross temporal properties. In additional tests, a 20 % change in note duty cycle, which exceeded the maximum CV_w reported for this property, successfully elicited recovery of aggression. This result indicates that sufficiently large differences in gross temporal properties might also contribute to the perceptual basis for discriminating between neighbors and strangers.

Recall that like songbirds, territorial male bullfrogs responded aggressively toward a neighbor's calls heard coming from a direction different from that of the neighbor's usual territory (Fig. 7.5b; Davis 1987). Thus, an important question was whether males could learn to associate a familiar voice recognition cue, such as fundamental frequency, with a particular location. Indeed, they could. When the original habituating stimulus was presented from a novel location during the discrimination phase, subjects resumed giving aggressive calls and aggressively approached the speaker in a novel location (Fig. 7.11c). A careful analysis of sounds presented from different locations indicated that acoustic differences imposed on the habituating stimulus as a result of propagation through slightly different microhabitats were too small to evoke renewed aggression from territorial males (Bee and Gerhardt 2001c). This assessment was based on comparing differences in the habituating stimuli broadcast and recorded from different locations to the larger changes imposed on novel stimuli in other playback tests that nevertheless failed to elicit recovery of aggressive responses. So after repeatedly hearing a call with a particular fundamental frequency coming from a particular direction, subjects learned to associate the call with its location of origin. The ability to make this association between a signal and a specific context as a result of repeated signal exposures probably serves as a basis for the bullfrog's ability to assign familiar neighbors to familiar locations.

7.4.3 Cognitive Basis of Neighbor Recognition in Bullfrogs

Social recognition in anurans, as exhibited by the dear enemy effect in bullfrogs, minimally requires that receivers possess mechanisms for learning, remembering, and categorizing vocal signals based on patterns of individual variation and distinctiveness. Moreover, learning about the specific spatial location from which signals usually originate represents the acquisition of a form of social knowledge about recognized individuals: "my neighbor belongs over *here*, not over *there*." In the previous section, I considered the decision rules governing the perceptual categorization of signals by territorial male bullfrogs as familiar versus unfamiliar. In this section, I review previous work aimed at elucidating mechanisms involved in learning about individually distinctive voice properties and their location of origin in this species.

Several authors have highlighted the potentially important role of habituation as a form of learning that could underlie the dear enemy effect (e.g., Peeke and Peeke

1973; Brooks and Falls 1975; Wiley and Wiley 1977; Peeke 1984; Petrinovich 1984; Shettleworth 2010). Habituation can be defined as “a behavioral response decrement that results from repeated stimulation and that does not involve sensory adaptation/sensory fatigue or motor fatigue” (p. 136 in Rankin et al. 2009). One key aspect of this definition is that habituation is *not* the same thing as sensory adaptation, the relatively short-term decrease in a neuron’s firing rate that results from the repeated or continuous presence of a stimulus. (In my experience, many people incorrectly equate habituation with sensory adaptation.) Nor does the response decrement result from simple “fatigue.” Rather, according to some accounts of perceptual learning, attention, and memory, habituation can be considered a cognitive process involved in the formation of persistent neuronal models of the physical properties of stimuli that can be associatively linked to other contextual stimuli, such as the environment in which stimuli are encountered (Wagner 1976, 1979, 1981; Schull 1979; Whitlow and Wagner 1984; Hall 1991; Siddle 1991; Cowan 1995; Tomsic et al. 2009). In this regard, habituation of a behavioral response reflects the operation of mechanisms that allow animals to form enduring representations of stimuli and to redirect their attention away from sources of irrelevant or unimportant stimulation. Empirical studies of aggression in fish (reviewed in Peeke and Peeke 1973; Peeke 1984), songbirds (reviewed in Petrinovich 1984), and frogs provide robust support for the general idea that habituation of aggression is involved in the dear enemy effect.

In a series of field playback studies, Bee (2001a, 2002, 2003a; Bee and Gerhardt 2001a, c, 2002) tested the hypothesis that the dear enemy effect in bullfrogs might result from long-term, stimulus-specific habituation. The goal of these experiments was to determine whether aggressive responses elicited by repeated exposure to the simulated calls of a new territorial neighbor exhibit the characteristics of habituation learning. In their seminal paper on habituation, Thompson and Spencer (1966) outlined nine parametric characteristics of changes in response due to habituation that occur with repeated stimulation. A recent review by Rankin et al. (2009) reaffirmed these nine characteristics of habituation learning and added a tenth (quoted here from that review):

1. *Repeated application of a stimulus results in a progressive decrease in some parameter of a response to an asymptotic level.*
2. *If the stimulus is withheld after response decrement, the response recovers at least partially over the observation time (“spontaneous recovery”).*
3. *After multiple series of stimulus repetitions and spontaneous recoveries, the response decrement becomes successively more rapid and/or more pronounced (this phenomenon can be called potentiation of habituation).*
4. *Other things being equal, more frequent stimulation results in more rapid and/or more pronounced response decrement, and more rapid spontaneous recovery (if the decrement has reached asymptotic levels).*
5. *Within a stimulus modality, the less intense the stimulus, the more rapid and/or more pronounced the behavioral response decrement. Very intense stimuli may yield no significant observable response decrement.*

6. *The effects of repeated stimulation may continue to accumulate even after the response has reached an asymptotic level (which may or may not be zero, or no response).*
7. *Within the same stimulus modality, the response decrement shows some stimulus specificity.*
8. *Presentation of a different stimulus results in an increase of the decremented response to the original stimulus. This phenomenon is termed "dishabituation."*
9. *Upon repeated application of the dishabituating stimulus, the amount of dishabituation produced decreases (this phenomenon can be called habituation of dishabituation).*
10. *Some stimulus repetition protocols may result in properties of the response decrement…that last hours, days or weeks.*

Eight of these 10 characteristics have been demonstrated to occur in the aggressive behavior of territorial bullfrogs in response to repeated playbacks of synthetic advertisement calls (Bee 2001a, 2002, 2003a; Bee and Gerhardt 2001a, c, 2002); the remaining two characteristics (#6 and #9) have not yet been examined. During the habituation phases of the experiments described in Sect. 7.4.2, males initially responded by producing advertisement calls and aggressive calls, and they approached the speaker with aggressive movements. In many instances, males continued to respond aggressively for several hours. With repeated stimulation, however, the numbers of aggressive calls, movements, and meters advanced toward the speaker declined, while the latency to respond aggressively to each stimulus increased (Fig. 7.12a). That is, aggressive responses decreased in magnitude to asymptotic levels (characteristic #1), and they typically did so following a negative exponential function, a response feature commonly observed in studies of habituation (Thompson and Spencer 1966). In addition, changes in the numbers of aggressive calls, movements, meters advanced toward the speaker, and response latency occurred at the same rates (Fig. 7.12a; Bee 2003a). This result strongly suggests that the neural mechanisms responsible for changes in different behaviors (e.g., vocalizations, phonotaxis) as a result of repeated stimulation operate at a locus in a receiver's brain that is common to different motor systems and that feeds forward into multiple motor circuits. As also discussed in Sect. 7.4.2, presentation of a novel stimulus in a subsequent discrimination phase, specifically one that differed from the habituating stimulus in either fundamental frequency or location, was able to elicit a recovery of habituated aggression. Thus, habituation was specific both to an individually distinctive voice property and the location from which the voice originated. These results, which unequivocally demonstrated the stimulus specificity of habituation (characteristic #7), were important for two reasons. First, they ruled out sensory adaptation/sensory fatigue and motor fatigue as explanations for the decrease in aggression that resulted from repeated playbacks during the habituation phase of the experiments. Second, these results were consistent with the hypothesis that stimulus-specific habituation could underlie the

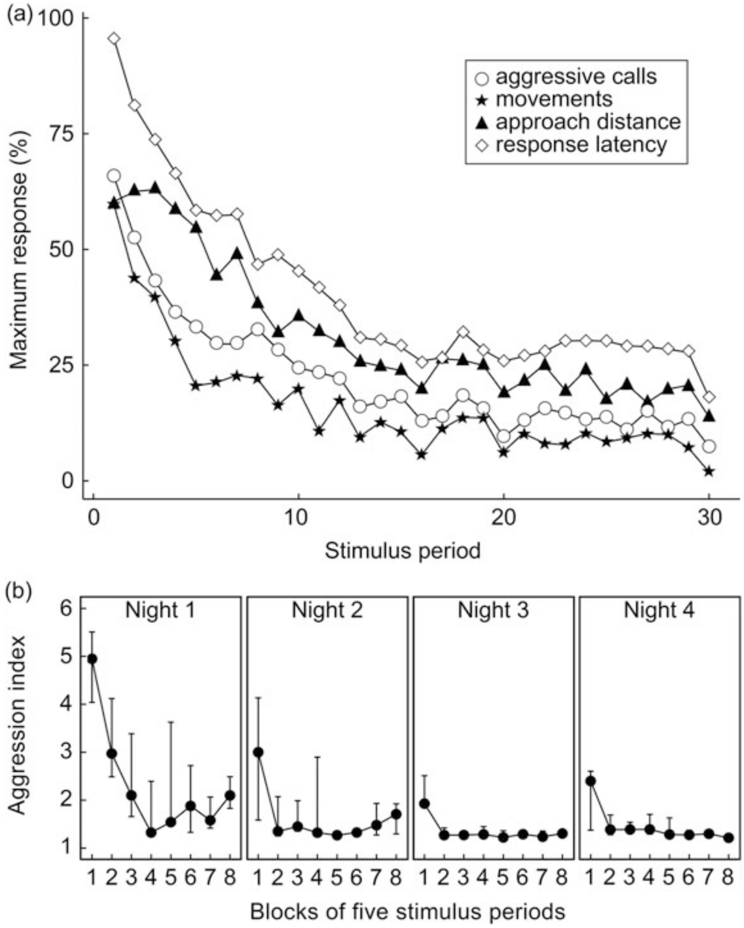


Fig. 7.12 Habituation of aggression in bullfrogs (*Rana catesbeiana*). (**a**) Within-night habituation. Over the course of 30 stimulus periods (3.75 h) within a single night of testing, repeated presentations of synthetic advertisement calls simulating the arrival of a new neighbor on an adjacent but previously unoccupied territory elicited marked decrements in aggressive responses. Shown here are the mean response magnitudes for the number of aggressive calls, the latency to the first aggressive call, the number of aggressive movements, and the maximum approach distance. Variables are expressed as a percentage of their maximum magnitude during any stimulus period (data redrawn from Bee 2003a). See legend for a description of symbols. (**b**) Long-term habituation. Shown here are the median and interquartile ranges of an aggressive index computed from a principal component analysis of the number of aggressive calls, the number of movements, and the maximum approach distance (data redrawn from Bee and Gerhardt 2001a). Subjects were tested over four consecutive nights. Within each night, subjects heard 40 stimulus periods (5 h) of a synthetic call broadcast from an adjacent but previously unoccupied territory. Aggressive responses habituated within each night. There was spontaneous recovery of responses at the start of each subsequent night, but the magnitude of spontaneous recovery declined across consecutive nights, and responses reached asymptotic levels more quickly after the first night of testing

stimulus specificity of reduced aggression that territory holders exhibit in response to hearing familiar neighbors calling from familiar locations.

Bee and Gerhardt (2001a) investigated spontaneous recovery (characteristic #2), potentiation of habituation (characteristic #3), and long-term habituation (characteristic #10) by repeatedly testing males over the course of multiple nights (Fig. 7.12b). On each of four consecutive nights, territorial males were presented with 5 h of repeated playbacks of the calls of a simulated new neighbor. One prediction was that aggressive responses would habituate within each night, but recover between nights in the absence of stimulation (i.e., spontaneous recovery). A second prediction was that, across nights, responses would tend to habituate more quickly within a night (i.e., potentiation of habituation). Together, support for these two predictions would be consistent with evidence for long-term habituation. As illustrated in Fig. 7.12b, habituated aggressive responses recovered spontaneously during the 19 intervening hours between the end of the 5-h habituation phase on one night and the beginning of the subsequent habituation phase on the next night. However, spontaneous recovery was not complete; that is, responses did not return to the levels where they started the preceding night. In addition, responses decreased to asymptotic levels more quickly across the four repeated nights of stimulus exposure (Fig. 7.12b). Subjects exhibited little in the way of aggressive responses by the fourth and final night of testing. In a discrimination phase that was tested on the fifth night, the long-term decrements in aggressive responsiveness that occurred during the preceding four nights recovered in response to changes in the fundamental frequency of the stimulus (Bee and Gerhardt 2001a). Taken together, these results demonstrated spontaneous recovery, potentiation of habituation, and long-term, stimulus-specific habituation. More importantly, the observation that subjects generally failed to respond to playbacks by the fourth night was consistent with earlier field observations reported by Haas (1977), according to which natural aggressive interactions between newly established neighbors tended to wane over the course of one to four nights.

In separate playback experiments, Bee (2001a) investigated whether the rate of habituation depended on stimulus repetition rate and stimulus intensity. These two features of the stimulus were manipulated using a factorial design. Stimulus repetition rate was manipulated across two levels (fast versus slow) by varying the interval between consecutive stimuli presented to different subjects during the habituation phase. Stimulus intensity was also manipulated across two levels (high versus low) by varying the sound pressure level of the repeated stimulus in tests of different subjects. Decreases in aggressive responsiveness were more rapid when the stimulus repetition rate was fast (characteristic #4) and when the stimulus intensity was low (characteristic #5). The slowest response decrements occurred when subjects heard stimuli that were presented with the combination of the slow repetition rate and high intensity.

Finally, Bee and Gerhardt (2001a) investigated dishabituation of aggression (characteristic #8). In the literature, the term "dishabituation" is often used incorrectly to refer to the recovery of habituation elicited in direct response to a novel stimulus. But dishabituation corresponds instead to the recovery of responses to subsequent presentations of the habituating stimulus following presentation of a "dishabituating stimulus" (Rankin et al. 2009). The dishabituating stimulus is typically a strong (and often novel) stimulus of some kind. Dishabituation reflects the operation of sensitization, which, along with habituation, is a key component of the dual-process theory of habituation (Groves and Thompson 1970). Bee and Gerhardt (2001a) showed that dishabituation of aggression in bullfrogs could be elicited by briefly presenting a dishabituating stimulus that differed from the habituating stimulus in fundamental frequency. The idea behind these tests was to first habituate a subject to a simulated new neighbor and then to simulate a brief intrusion by a different individual. Compared with subjects in a control group, males that were exposed to the dishabituating intruder stimulus were more likely to respond aggressively to subsequent presentations of the original habituating stimulus. The interpretation of this result, according to the dual-process theory of habituation, is that the brief intrusion by a novel opponent caused subjects to become sensitized, which in turn caused an increase in the magnitude of the formerly habituated aggression exhibited toward the familiar neighbor. These and other results (see Bee 2001a) provided support for the notion that two processes—habituation and sensitization—determine the magnitude of aggressive responses elicited by repeated stimuli that simulate calling males.

Together, results from these studies of territorial aggression in bullfrogs support the hypothesis that learning—specifically learning in the form of long-term, stimulus-specific habituation—occurs in response to hearing the repeated signals of a new neighbor and can produce a dear enemy effect (see also Owen and Perrill 1998; Bee and Schachtman 2000). As a result of learning, territory residents encode information about a neighbor's individually distinctive voice properties and associate them with a particular location. Notably, the patterns of changes in aggression observed in bullfrogs were strikingly similar to those previously demonstrated to occur in territorial fish (reviewed in Peeke and Peeke 1973; Peeke 1984) and songbirds (reviewed in Petrinovich 1984). The hypothesis that habituation can generate a dear enemy effect is now well accepted (Shettleworth 2010; Wiley 2013).

As I have argued before (Bee and Gerhardt 2001a), attributing a role for habituation in the dear enemy effect in no way excludes the operation of potentially more complex forms of cognitive processing. In fact, we should expect more complex processes to be involved in some animals (Bee 2006). Male songbirds, for example, exhibit a dear enemy effect, and they exhibit long-term, stimulus-specific, and location-specific habituation of territorial aggression. But some songbird species, such as the song sparrow (*Melospiza melodia*), also possess detailed social knowledge pertaining to the song repertoires of their neighbors (Beecher et al. 1996). Another songbird, the rufous-sided towhee (*Pipilo erythrophthalmus*), has been shown to learn to recognize the songs of neighboring males that actually sang imitations of a different species' songs (Richards 1979). Some have argued that such behaviors rule out habituation and require some form of associative learning (Richards 1979; Wiley 2013). However, the distinction between habituation and associative learning can be far less clear than some might wish. For example, there are explanations of habituation based on associative learning, such as Wagner's associative theory of habituation, in which long-term habituation involves the formation of an association between the habituating stimulus and the specific context in which it is presented (Wagner 1976, 1979, 1981; Whitlow and Wagner 1984). This theory has received strong support in studies of the natural behaviors of some animals (e.g., Tomsic et al. 2009). Elsewhere, I have argued that the location from which a neighbor's signals originate can function as an important contextual cue for recognition (Bee 2001b; Bee and Gerhardt 2002). The key take-home point here is that drawing a stark and strict boundary between habituation versus associative learning creates a false dichotomy that is not only inconsistent with some views of long-term habituation but that also potentially prevents forward progress on understanding the diversity of cognitive mechanisms underlying social recognition. Mere distinctions between labels might tell us less than we think about underlying mechanisms at the cellular and network levels.

An important question that has not yet been addressed in any depth in anurans concerns the social knowledge that territory residents acquire about their neighbors. I would argue that the contextual information related to the position from which a neighbor usually calls constitutes a form of social knowledge that can be acquired by repeatedly hearing a neighbor's calls originate from a particular location. We should not be surprised if territorial anurans can acquire other social knowledge about their neighbors in addition to their usual location. For example, neighboring male bullfrogs interact extensively as they sort out common territorial borders (Emlen 1968, 1976; Wiewandt 1969; Howard 1978a; Ryan 1980). In addition to bouts of calling and counter calling, these interactions sometimes escalate into intense physical fights lasting an hour or more (Fig. 7.4). It seems plausible that male bullfrogs could acquire information about the size and fighting ability of their new neighbors through these vocal and physical interactions. This information would likely influence a male's propensity to interact aggressively with his neighbors. Territory residents might indeed be "fighting to learn" (Getty 1989).

7.4.4 Section Summary

The studies of North American bullfrogs reviewed in Sect. 7.4, and earlier in Sect. 7.3.1, have demonstrated a number of key findings regarding social recognition in this species. First, males defend long-term territories that contain a valuable breeding resource in the form of oviposition sites. Second, territorial males behaviorally discriminate between neighbors and strangers in ways consistent with the dear enemy effect, and they do so based on individual differences in vocalizations. Moreover, territory residents are able to associate familiar neighbors with familiar locations. Third, most properties of bullfrog advertisement calls are individually distinctive, though fundamental frequency and correlated spectral properties contribute most toward individual distinctiveness. Fourth, playback tests based on the habituation-discrimination paradigm have shown that males discriminate between familiar and unfamiliar calls based on differences in fundamental frequency as small as 5–10 %, which maps closely onto expectations about JMDs based on patterns of within-individual and among-individual differences. Finally, decreases in aggression in response to simulated new territorial neighbors exhibit the characteristics of response habituation and are both long lasting and specific to individually distinctive voice properties and location. These results indicate long-term habituation is likely an important learning mechanism underlying neighbor recognition and the dear enemy effect.

7.5 Summary and Future Directions

We will never know whether Linnaeus would have believed that anurans can learn to recognize the "harsh call" of a competitive rival. But this appears to be the case, at least for some species. Social recognition, as evidenced by the dear enemy effect, may play important roles in mediating aggressive interactions between territorial male anurans. So far, this form of social recognition has been reported to occur in four ranid species (bullfrogs, agile frogs, concave-eared torrent frogs, and olive frogs) as well as in one dendrobatid (golden rocket frogs) but not in another (strawberry poison frogs). Hence, there appears to have been convergent evolution, perhaps multiple times (Fig. 7.1), in the vocally mediated social behaviors of territorial anurans and songbirds. However, one thing I hope this review makes clear is that not all studies of the dear enemy phenomenon should inspire equal confidence in the presence or absence of this behavior in anurans. The interpretation of results from some studies are complicated by various methodological problems. Elsewhere, my colleagues and I have provided an in-depth treatment of best practices that future researchers can follow to avoid these difficulties of interpretation (Bee et al. 2016). In the remainder of this section, I provide a roadmap for making forward progress toward understanding the evolution and mechanisms of social recognition in anurans.

7.5.1 Evolution of Social Recognition and the Dear Enemy Effect

At present, we lack a clear understanding of the ecological and social factors that select for dear enemy effects in territorial systems. What is clear is that simply having territorial neighbors is not enough (Bee 2003b). Temeles (1994) has argued, based on an extensive literature review, that territorial neighbors are more likely to display reduced levels of aggression toward each other when a neighbor poses less of a threat than a non-territorial floater (i.e., a stranger). In his analysis, the dear enemy effect was more frequently found in species in which individuals defend breeding or multipurpose territories compared with species that defend primarily feeding territories. A proposed explanation for this pattern is that when territory holders defend breeding or multipurpose territories, neighbors may occasionally steal mating opportunities, but they already defend a territory, whereas a stranger may attempt to steal not just the occasional mate but also the defended resources required for a successful breeding attempt in the first place (Temeles 1994). Hence, strangers are thought to pose the greater relative threat when territorial animals defend breeding or multipurpose territories.

Anurans represent one of the best taxonomic groups in which to test Temeles' (1994) ideas and, more broadly, in which to investigate the ecological and social factors that favor the evolution of the dear enemy effect and social recognition. This is because frogs and toads exhibit a diversity of mating systems that are not only evolutionarily labile but can also vary within a species depending on resource availability in time and space (Wells 1977, 2007). Although we lack sufficient natural history data to be certain, it is probably not uncommon for even closely related species (e.g., within the same genus) to exhibit different mating systems, such as in North American *Rana* (Wells 1977, 2007). Moreover, territorial systems in which males defend breeding or multipurpose territories have almost certainly had multiple, independent evolutionary origins in anurans. A major goal for future research should be to capitalize on this diversity to test hypotheses about the ecological and social factors that favor social recognition of territorial neighbors.

Going forward, two lessons are to be learned from the case studies reviewed in Sect. 7.3. First, preliminary evidence is at least partially consistent with the notion that dear enemy effects are observed in species that defend breeding or multipurpose territories. In bullfrogs, golden rocket frogs, and olive frogs, males defend territories in which females oviposit; hence, territory residents in these species defend a key breeding resource, and they also respond less aggressively toward neighbors than strangers. Data on strawberry poison frogs are more equivocal. There is, so far, little evidence that territory holders behaviorally discriminate between neighbors and strangers. Although early reports described territories in this species as containing oviposition sites and tadpole-rearing sites (two breeding resources), these findings have been questioned more recently. And this brings us to the second lesson to be learned: accurate data on a species' natural history and behavior are *critical* for placing evidence for and against the dear enemy effect in proper ecological and social contexts. Studies of agile frogs and concave-eared torrent frogs illustrate this lesson well. Without basic natural history data, results from empirical studies of the dear enemy effect, whether positive or negative, are exceedingly difficult to interpret, and they add little to our understanding of how the phenomenon evolves. In the future, we will need to investigate social recognition in anurans using a phylogenetic framework (Fig. 7.1) that is informed by rigorous data on basic natural history and behavior and that attempts to describe the nature of defended resources and to quantify the relative threats posed by neighbor and strangers.

A second future goal of work on social recognition in anurans should be to identify its evolutionary function in terms of benefits and costs. It is common to read in the literature that the benefit of the dear enemy effect is in allowing territory holders to avoid repeated aggressive interactions that could otherwise result in increased energy expenditure, increased risk of injury, increased predation risk (e.g., due to decreased vigilance), and decreased mate attraction. Thus, the benefit of dear enemies may often derive from cost savings. Yet few studies have attempted to quantify these costs. More importantly, any savings accrued through avoiding repeated aggressive interactions with neighbors must be weighed against the energetic costs of developing and maintaining the perceptual and cognitive mechanisms necessary for recognition. Little effort has been made so far to quantify these costs of recognition in any species. Quantifying these costs, of course, requires that we identify the mechanisms of social recognition and the dear enemy effect.

7.5.2 Mechanisms of Social Recognition and the Dear Enemy Effect

Understanding proximate mechanisms is absolutely critical to understanding behavioral adaptations. Substantial future progress stands to be made in identifying proximate mechanisms for social recognition in anurans by targeting the key components of recognition systems described in Sect. 7.2, which highlight adaptations of both signalers and receivers.

According to the signature adaptation hypothesis, factors that favor recognition can select for the production of more individually distinctive signals by signalers. Robust support for this hypothesis has been found in studies of swallows and wasps (e.g., Beecher 1988, 1991; Tibbetts and Dale 2007). For social recognition in anurans, the signature adaptation hypothesis makes the following prediction: males in territorial species that exhibit the dear enemy effect should produce more individually distinctive calls than males of closely related but non-territorial species that do not exhibit the phenomenon. Testing this hypothesis in anurans will require detailed acoustic and statistical analyses of signals performed in a comparative, phylogenetic framework. Obviously, such a framework would have to be informed by rigorous studies of natural history and behavior. Beecher (1989b) has developed an approach that provides a means of quantifying and computing across species the information content of signals relevant to individual identity. Future comparative studies of the signature adaptation hypothesis using Beecher's (1989b) methods to analyze anuran vocalizations seem easily within grasp.

Adaptations related to the perceptual and cognitive processes involved in recognition by receivers should also be investigated at a mechanistic level. In anurans, an important future goal should be to investigate the evolutionary match between the patterns of individual variation in signals and the perceptual basis of recognition, specifically the magnitude of JMDs across signal properties. Addressing this core issue has potential to elucidate the boundaries of learned categories of individuals. To what extent do natural and sexual selection favor drawing categorical boundaries using the most individually distinctive cues? Comparative studies of species that are similar in recognizing neighbors but different in terms of the signal properties that contribute most to individual distinctiveness (e.g., bullfrogs versus golden rocket frogs) would be worthwhile. Along similar lines of inquiry, what are the specificity and multiplicity (sensu Wiley 2013) of learned social categories in anurans? Can territorial males learn to recognize multiple individual neighbors? Davis's (1987) data showing that bullfrogs associate neighbors with particular locations is suggestive, but inconclusive. A related question concerns how social categories of individuals are learned and what kinds of social knowledge are cognitively bound to representations of familiar signals. Is a form of habituation that is long lasting, stimulus specific, and context specific the whole story in anurans? Can frogs form cross-modal representation of individuals (Proops et al. 2009; Chap. 8)? How does the learned recognition of individual identity cues map onto the unlearned recognition of species identity cues (see Nelson 1989 for related work on a similar question in songbirds)? At a neuronal level, it will be important to determine how and where in the brain the information necessary for social recognition is coded in neural responses. Studies of immediate early gene expression and electrophysiological recordings might provide some answers to questions about neural mechanisms, as they have in studies of individual recognition in songbirds (e.g., Mello et al. 1995; Gentner 2004).

7.5.3 Conclusion

Given the number of anuran species, and the diversity of social and reproductive behaviors that has evolved in this group, anurans represent an important vertebrate taxon in which to investigate the mechanisms and evolution of vocally mediated social recognition. The data reviewed here suggest a new, concerted research agenda on social recognition in anurans would be worthwhile. Going forward, these efforts will require robust descriptive studies of natural history, rigorous field experiments that investigate behavioral discrimination of learned social categories, careful consideration of phylogeny, and the integrative use of various empirical tools and conceptual approaches to study the major components of recognition systems.

Acknowledgments The author is grateful to the National Science Foundation, the Smithsonian Tropical Research Institute, Sigma Xi, and the University of Minnesota for funding various portions of the original research reviewed in this chapter.

References

AmphibiaWeb (2016) Information on amphibian biology and conservation. AmphibiaWeb, Berkeley. http://amphibiaweb.org/. Last accessed 12 Feb 2016

Arch VS, Burmeister SS, Feng AS, Shen JX, Narins PM (2011) Ultrasound-evoked immediate early gene expression in the brainstem of the Chinese torrent frog, *Odorrana tormota*. J Comp Physiol A 197(6):667–675

Baugh JR, Forester DC (1994) Prior residence effect in the dart-poison frog, *Dendrobates pumilio*. Behaviour 131(3/4):207–224

Bee MA (2001a) Habituation and sensitization of aggression in bullfrogs (*Rana catesbeiana*): testing the dual-process theory of habituation. J Comp Psychol 115(3):307–316

Bee MA (2001b) Vocally mediated neighbor recognition in North American bullfrogs, *Rana catesbeiana*: identification, perception, and learning. Ph.D. Dissertation, University of Missouri, Columbia

Bee MA (2002) Territorial male bullfrogs (*Rana catesbeiana*) do not assess fighting ability based on size-related variation in acoustic signals. Behav Ecol 13(1):109–124

Bee MA (2003a) Experience-based plasticity of acoustically evoked aggression in a territorial frog. J Comp Physiol A 189(6):485–496

Bee MA (2003b) A test of the 'dear enemy effect' in the strawberry dart-poison frog (*Dendrobates pumilio*). Behav Ecol Sociobiol 54(6):601–610

Bee MA (2004) Within-individual variation in bullfrog vocalizations: implications for a vocally mediated social recognition system. J Acoust Soc Am 116(6):3770–3781

Bee MA (2006) Individual recognition in animal species. In: Naguib M (ed) Encyclopedia of language and linguistics, 2nd edn. Elsevier, New York, pp. 617–626

Bee MA, Gerhardt HC (2001a) Habituation as a mechanism of reduced aggression between neighboring territorial male bullfrogs (*Rana catesbeiana*). J Comp Psychol 115(1):68–82

Bee MA, Gerhardt HC (2001b) Neighbour-stranger discrimination by territorial male bullfrogs (*Rana catesbeiana*): I. acoustic basis. Anim Behav 62:1129–1140

Bee MA, Gerhardt HC (2001c) Neighbour-stranger discrimination by territorial male bullfrogs (*Rana catesbeiana*): II. perceptual basis. Anim Behav 62:1141–1150

Bee MA, Gerhardt HC (2002) Individual voice recognition in a territorial frog (*Rana catesbeiana*). Proc R Soc Lond B Biol Sci 269(1499):1443–1448

Bee MA, Schachtman TR (2000) Is habituation a mechanism for neighbor recognition in green frogs? Behav Ecol Sociobiol 48(2):165–168

Bee MA, Kozich CE, Blackwell KJ, Gerhardt HC (2001) Individual variation in advertisement calls of territorial male green frogs, Rana clamitans: implications for individual discrimination. Ethology 107(1):65–84

Bee MA, Reichert MS, Tumulty JP (2016) Assessment and recognition of competitive rivals in anurans. Adv Stud Behav 48:161–249

Beecher MD (1982) Signature systems and kin recognition. Am Zool 22(3):477–490

Beecher MD (1988) Kin recognition in birds. Behav Genet 18(4):465–482

Beecher MD (1989a) Evolution of parent-offspring recognition in swallows. In: Dewsbury DA (ed) Contemporary issues in comparative psychology. Sinauer, Sunderland, pp. 360–380

Beecher MD (1989b) Signalling systems for individual recognition: an information theory approach. Anim Behav 38:248–261

Beecher MD (1991) Successes and failures of parent-offspring recognition systems in animals. In: Hepper PG (ed) Kin recognition. Cambridge University Press, Cambridge, pp. 94–124

Beecher MD, Loesche P, Stoddard PK, Medvin MB (1989) Individual recognition by voice in swallows: signal or perceptual adaptation. In: Dooling RJ, Hulse SH (eds) The comparative psychology of audition: perceiving complex sounds. Lawrence Erlbaum Associates, Hillsdale, pp. 277–292

Beecher MD, Stoddard PK, Campbell ES, Horning CL (1996) Repertoire matching between neighbouring song sparrows. Anim Behav 51(4):917–923

Beer CG (1970) Individual recognition of voice in the social behavior of birds. Adv Study Behav 3:27–74

Boake CRB (1989) Repeatability: its role in evolutionary studies of mating behavior. Evol Ecol 3 (2):173–182

Boatright-Horowitz SS, Cheney CA, Simmons AM (1999) Atmospheric and underwater propagation of bullfrog vocalizations. Bioacoustics 9:257–280

Bossuyt F, Roelants K (2009) Frogs and toads (Anura). In: Hedges SB, Kumar S (eds) The timetree of life. Oxford University Press, New York, pp. 357–364

Bourne GR (1992) Lekking behavior in the neotropical frog *Ololygon rubra*. Behav Ecol Sociobiol 31(3):173–180

Bourne GR, Collins AC, Holder AM, McCarthy CL (2001) Vocal communication and reproductive behavior of the frog *Colostethus beebei* in Guyana. J Herpetol 35(2):272–281

Bradbury JW, Vehrencamp SL (2011) Principles of animal communication, 2nd edn. Sinauer Associates, Sunderland

Brenowitz EA, Rose GJ (1994) Behavioral plasticity mediates aggression in choruses of the Pacific treefrog. Anim Behav 47(3):633–641

Briefer E, Aubin T, Lehongre K, Rybak F (2008) How to identify dear enemies: the group signature in the complex song of the skylark *Alauda arvensis*. J Exp Biol 211(3):317–326

Brooks RJ, Falls JB (1975) Individual recognition by song in white-throated sparrows. I discrimination of songs of neighbors and strangers. Can J Zool 53(7):879–888

Brown JL, Morales V, Summers K (2010) A key trait drove the evolution of biparental care and monogamy in an amphibian. Am Nat 175(4):436–446

Bunnell P (1973) Vocalizations in the territorial behavior of the frog *Dendrobates pumilio*. Copeia:277–284

Catchpole CK (1989) Pseudoreplication and external validity: playback experiments in avian bioacoustics. Trends Ecol Evol 4(10):286–287

Chuang M-F, Bee MA, Kam Y-C (2013) Short amplexus duration in a territorial anuran: a possible adaptation in response to male-male competition. PLoS One 8(12):e83116

Chuang M-F, Kam Y-C, Bee MA (2016) Quantitative description of the vocal repertoire of the territorial olive frog *Babina adenopleura* from Taiwan. Bioacoustics 25(1):1–18

Chuang M-F, Kam Y-C, Bee MA (2017) Vocally-mediated discrimination between neighbors and strangers in a territorial frog, *Babina adenopleura*. Anim Behav 123:217–228

Chuaynkern Y, Ohler A, Inthara C, Duengkae P, Makchai S, Salangsingha N (2010) A revision of species in the subgenus *Nidirana* Dubois, 1992, with special attention to the identity of specimens allocated to *Rana adenopleura* Boulenger, 1909, and *Rana chapaensis* (Bourret, 1937)(Amphibia: Anura: Ranidae) from Thailand and Laos. Raffles Bull Zool 58(2):291–310

Colgan PW (1983) Comparative social recognition. Wiley, New York

Cowan N (1995) Attention and memory: an integrated framework, Oxford Psychology Series, vol 26. Calrendon Press, Oxford

Crothers LR, Cummings ME (2015) A multifunctional warning signal behaves as an agonistic status signal in a poison frog. Behav Ecol 26(2):560–568

Davis MS (1987) Acoustically mediated neighbor recognition in the North American bullfrog, *Rana catesbeiana*. Behav Ecol Sociobiol 21(3):185–190

Donnelly MA (1989a) Demographic effects of reproductive resource supplementation in a territorial frog, *Dendrobates pumilio*. Ecol Monogr 59(3):207–221

Donnelly MA (1989b) Effects of reproductive resource supplementation on space-use patterns in *Dendrobates pumilio*. Oecologia 81(2):212–218

Donnelly MA (1989c) Phenology and age structure of *Dendrobates pumilio* in northeastern Costa Rica. J Herpetol 23(4):362–367

Dyson ML, Reichert MS, Halliday TR (2013) Contests in amphibians. In: Hardy ICW, Briffa M (eds) Animal contests. Cambridge University Press, Cambridge, pp. 228–257

Emlen ST (1968) Territoriality in the bullfrog, *Rana catesbeiana*. Copeia:240–243

Emlen ST (1976) Lek organization and mating strategies in the bullfrog. Behav Ecol Sociobiol 1 (3):283–313

Falls JB (1982) Individual recognition by sounds in birds. In: Kroodsma DE, Miller EH (eds) Acoustic communication in birds. Academic, New York, pp. 237–278

Feng AS, Narins PM (2008) Ultrasonic communication in concave-eared torrent frogs (*Amolops tormotus*). J Comp Physiol A 194(2):159–167

Feng AS, Narins PM, Xu CH (2002) Vocal acrobatics in a Chinese frog, *Amolops tormotus*. Naturwissenschaften 89(8):352–356

Feng AS, Narins PM, Xu CH, Lin WY, Yu ZL, Qiu Q, Xu ZM, Shen JX (2006) Ultrasonic communication in frogs. Nature 440(7082):333–336

Feng AS, Arch VS, Yu ZL, Yu XJ, Xu ZM, Shen JX (2009a) Neighbor-stranger discrimination in concave-eared torrent frogs, *Odorrana tormota*. Ethology 115(9):851–856

Feng AS, Riede T, Arch VS, Yu ZL, Xu ZM, Yu XJ, Shen JX (2009b) Diversity of the vocal signals of concave-eared torrent frogs (*Odorrana tormota*): Evidence for individual signatures. Ethology 115(11):1015–1028

Fisher JB (1954) Evolution and bird sociality. In: Huxely J, Hardy AC, Ford EB (eds) Evolution as a process. Allen & Unwin, London, pp. 71–83

Frost DR (2015) Amphibian species of the world: an online reference. Version 6.0 (12 February 2016). Electronic Database accessible at http://research.amnh.org/herpetology/amphibia/index.html. American Museum of Natural History, New York

Frost DR, Grant T, Faivovich J, Bain RH, Haas A, Haddad CFB, De Sa RO, Channing A, Wilkinson M, Donnellan SC, Raxworthy CJ, Campbell JA, Blotto BL, Moler P, Drewes RC, Nussbaum RA, Lynch JD, Green DM, Wheeler WC (2006) The amphibian tree of life. Bull Am Mus Nat Hist 297:8–370

Gasser H, Amézquita A, Hödl W (2009) Who is calling? Intraspecific call variation in the aromobatid frog *Allobates femoralis*. Ethology 115(6):596–607

Gentner TQ (2004) Neural systems for individual song recognition in adult birds. In: Zeigler HP, Marler P (eds) Behavioral neurobiology of birdsong, Annals of the new york academy of sciences, vol 1016. New York Academic Sciences, New York, pp. 282–302

Gerhardt HC, Huber F (2002) Acoustic communication in insects and anurans: common problems and diverse solutions. Chicago University Press, Chicago

Gerhardt HC, Schwartz JJ (2001) Auditory tuning, frequency preferences and mate choice in anurans. In: Ryan MJ (ed) Anuran communication. Smithsonian Institution Press, Washington, DC, pp. 73–85

Gerhardt HC, Roberts JD, Bee MA, Schwartz JJ (2000) Call matching in the quacking frog (*Crinia georgiana*). Behav Ecol Sociobiol 48(3):243–251

Getty T (1989) Are dear enemies in a war of attrition? Anim Behav 37:337–339

Gridi-Papp M, Feng AS, Shen JX, Yu ZL, Rosowski JJ, Narins PM (2008) Active control of ultrasonic hearing in frogs. Proc Natl Acad Sci USA 105(31):11014–11019

Groves PM, Thompson RF (1970) Habituation: a dual-process theory. Psychol Rev 77(5):419–450

Haas JD (1976) Individual differences in the vocalizations of the bullfrog, *Rana catesbeiana*. Herpetol Rev 7:86

Haas JD (1977) Individual differences in the vocalization of bullfrogs, *Rana catesbeiana*. Ph.D. Dissertation, Tufts University, Medford

Haase A, Pröhl H (2002) Female activity patterns and aggressiveness in the strawberry poison frog *Dendrobates pumilio* (Anura: Dendrobatidae). Amphibia-Reptilia 23(2):129–140

Hall G (1991) Perceptual and associative learning, Oxford psychology series, vol 18. Clarendon Press, Oxford

Halpin ZT (1986) Individual odors among mammals: origins and functions. Adv Study Behav 16:39–70

Hedges SB, Kumar S (2009) The timetree of life. Oxford University Press, Oxford

Howard RD (1978a) Evolution of mating strategies in bullfrogs, *Rana catesbeiana*. Evolution 32 (4):850–871

Howard RD (1978b) Influence of male-defended oviposition sites on early embryo mortality in bullfrogs. Ecology 59(4):789–798

Howard RD (1979) Estimating reproductive success in natural populations. Am Nat 114 (2):221–231

Howard RD (1981) Sexual dimorphism in bullfrogs. Ecology 62(2):303–310

Howard RD (1983) Sexual selection and variation in reproductive success in a long-lived organism. Am Nat 122(3):301–325

Howard RD (1984) Alternative mating behaviors of young male bullfrogs. Am Zool 24 (2):397–406

Kitchell K Jr, Dundee HA (1994) A trilogy on the herpetology of Linnaeus's *Systema Naturae X*. Smithsonian Herpetol Inf Serv 100:1–61

Kroodsma DE (1989a) Inappropriate experimental designs impede progress in bioacoustic research: a reply. Anim Behav 38:717–719

Kroodsma DE (1989b) Suggested experimental designs for song playbacks. Anim Behav 37 (4):600–609

Kroodsma DE (1990) Using appropriate experimental designs for intended hypotheses in 'song' playbacks, with examples for testing effects of song repertoire sizes. Anim Behav 40 (6):1138–1150

Kroodsma DE, Byers BE, Goodale E, Johnson S, Liu WC (2001) Pseudoreplication in playback experiments, revisited a decade later. Anim Behav 61:1029–1033

Lambrechts MM, Dhondt AA (1995) Individual voice discrimination in birds. In: Power DM (ed) Current ornithology. Plenum, New York, pp. 115–139

Lesbarrères D, Lodé T (2002) Variations in male calls and responses to an unfamiliar advertisement call in a territorial breeding anuran, *Rana dalmatina*: evidence for a "dear enemy" effect. Ethol Ecol Evol 14(4):287–295

Lesbarrères D, Merila J, Lodé T (2008) Male breeding success is predicted by call frequency in a territorial species, the agile frog (*Rana dalmatina*). Can J Zool 86(11):1273–1279

Lessells CM, Boag PT (1987) Unrepeatable repeatabilities: a common mistake. Auk 104 (1):116–121

Linnaeus C (1758) Systema naturae per regna tria naturae: secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis, 10th edn. Laurentius Salvius, Stockholm

Lodé T (2009) Mating strategies and monogamy in a territorial breeding anuran, *Rana dalmatina*: a result of sexual conflict? Alytes (Paris) 27(2):37–48

Lodé T, Lesbarrères D (2004) Multiple paternity in *Rana dalmatina*, a monogamous territorial breeding anuran. Naturwissenschaften 91(1):44–47

Lodé T, Holveck MJ, Lesbarrères D (2005) Asynchronous arrival pattern, operational sex ratio and occurrence of multiple paternities in a territorial breeding anuran, *Rana dalmatina*. Biol J Linn Soc 86(2):191–200

Maan ME, Cummings ME (2012) Poison frog colors are honest signals of toxicity, particularly for bird predators. Am Nat 179(1):E1–E14

Marshall VT, Humfeld SC, Bee MA (2003) Plasticity of aggressive signalling and its evolution in male spring peepers, *Pseudacris crucifer*. Anim Behav 65:1223–1234

McGregor PK, Catchpole CK, Dabelsteen T, Falls JB, Fusani L, Gerhardt HC, Gilbert F, Horn A, Klump GM, Kroodsma DE, Lambrechts MM, McComb KE, Nelson DA, Pepperberg IM, Ratcliffe L, Searcy WA, Weary DM (1992) Design of playback experiments: the Thornbridge Hall NATO ARW consensus. In: McGregor PK (ed) Playback and studies of animal communication. Plenum Press, New York, pp. 1–9

McVey ME, Zahary RG, Perry D, Macdougal J (1981) Territoriality and homing behaviour in the poison dart frog (*Dendrobates pumilio*). Copeia 1:1–8

Mello C, Nottebohm F, Clayton D (1995) Repeated exposure to one song leads to a rapid and persistent decline in an immediate-early genes response to that song in zebra finch telencephalon. J Neurosci 15(10):6919–6925

Meuche I, Linsenmair KE, Pröhl H (2012) Intrasexual competition, territoriality and acoustic communication in male strawberry poison frogs (*Oophaga pumilio*). Behav Ecol Sociobiol 66 (4):613–621

Miller CT, Bee MA (2012) Receiver psychology turns 20: is it time for a broader approach? Anim Behav 83(2):331–343

Narins PM, Feng AS, Lin WY, Schnitzler HU, Denzinger A, Suthers RA, Xu CH (2004) Old World frog and bird vocalizations contain prominent ultrasonic harmonics. J Acoust Soc Am 115(2):910–913

Narins PM, Feng AS, Fay RR, Popper AN (2007) Hearing and sound communication in amphibians, Springer handbook of auditory research, vol 28. Springer, New York

Nelson DA (1989) Song frequency as a cue for recognition of species and individuals in the field sparrow (*Spizella pusilla*). J Comp Psychol 103(2):171–176

Nelson DA, Marler P (1990) The perception of birdsong and an ecological concept of signal space. In: Berkley MA, Stebbins WC (eds) Comparative perception: volume II. Wiley, New York, pp. 443–478

Owen PC, Perrill SA (1998) Habituation in the green frog, *Rana clamitans*. Behav Ecol Sociobiol 44(3):209–213

Peeke HVS (1984) Habituation and the maintenance of territorial boundaries. In: Peeke HVS, Petrinovich L (eds) Habituation, sensitization, and behavior. Academic, New York, pp. 393–421

Peeke HVS, Peeke SC (1973) Habituation in fish with special reference to intraspecific aggressive behavior. In: HVS P, MJ H (eds) Habituation: I. behavioral studies. Academic, New York, pp. 59–85

Petrinovich L (1984) A two-factor dual-process theory of habituation and sensitization. In: Peeke HVS, Petrinovich L (eds) Habituation, sensitization, and behavior. Academic, New York, pp. 17–55

Pettitt BA (2012) Paternal effects in relation to acoustically mediated mate choice in a neotropical frog. Ph.D. Thesis, University of Minnesota, Minneapolis

Pettitt BA, Bourne GR, Bee MA (2012) Quantitative acoustic analysis of the vocal repertoire of the golden rocket frog (*Anomaloglossus beebei*). J Acoust Soc Am 131:4811–4820

Pettitt BA, Bourne GR, Bee MA (2013) Advertisement call variation in the golden rocket frog (*Anomaloglossus beebei*): evidence for individual distinctiveness. Ethology 119:244–256

Pröhl H (1997a) Reproductive pattern in *Dendrobates pumilio* (Anura: Dentrobatidae). Rev Biol Trop 45(4):1669–1674

Pröhl H (1997b) Territorial behaviour of the strawberry poison-dart frog, *Dendrobates pumilio*. Amphibia-Reptilia 18(4):437–442

Pröhl H (2002) Population differences in female resource abundance, adult sex ratio, and male mating success in *Dendrobates pumilio*. Behav Ecol 13(2):175–181

Pröhl H (2003) Variation in male calling behaviour and relation to male mating success in the strawberry poison frog (*Dendrobates pumilio*). Ethology 109(4):273–290

Pröhl H (2005a) Clutch loss affects the operational sex ratio in the strawberry poison frog *Dendrobates pumilio*. Behav Ecol Sociobiol 58(3):310–315

Pröhl H (2005b) Territorial behavior in dendrobatid frogs. J Herpetol 39(3):354–365

Pröhl H, Berke O (2001) Spatial distributions of male and female strawberry poison frogs and their relation to female reproductive resources. Oecologia 129(4):534–542

Pröhl H, Hödl W (1999) Parental investment, potential reproductive rates, and mating system in the strawberry dart-poison frog, *Dendrobates pumilio*. Behav Ecol Sociobiol 46(4):215–220

Proops L, McComb K, Reby D (2009) Cross-modal individual recognition in domestic horses (*Equus caballus*). Proc Natl Acad Sci USA 106(3):947–951

Pyron RA, Wiens JJ (2011) A large-scale phylogeny of Amphibia including over 2800 species, and a revised classification of extant frogs, salamanders, and caecilians. Mol Phylogenet Evol 61 (2):543–583

Rankin CH, Abrams T, Barry RJ, Bhatnagar S, Clayton DF, Colombo J, Coppola G, Geyer MA, Glanzman DL, Marsland S (2009) Habituation revisited: an updated and revised description of the behavioral characteristics of habituation. Neurobiol Learn Mem 92(2):135–138

Reichert MS, Gerhardt HC (2012) Trade-offs and upper limits to signal performance during close-range vocal competition in gray tree frogs *Hyla versicolor*. Am Nat 180(4):425–437

Richards DG (1979) Recognition of neighbors by associative learning in rufous-sided towhees. The Auk 96:688–693

Richards DG, Wiley RH (1980) Reverberations and amplitude fluctuations in the propagation of sound in a forest: implications for animal communication. Am Nat 115(3):381–399

Robertson JGM (1984) Acoustic spacing by breeding males of *Uperoleia rugosa* (Anura, Leptodactylidae). Z Tierpsychol 64(3–4):283–297

Rose GJ, Brenowitz EA (1991) Aggressive thresholds of male Pacific treefrogs for advertisement calls vary with amplitude of neighbors' calls. Ethology 89(3):244–252

Ryan MJ (1980) The reproductive behavior of the bullfrog (*Rana catesbeiana*). Copeia 1980 (1):108–114

Ryan MJ, Keddy-Hector A (1992) Directional patterns of female mate choice and the role of sensory biases. Am Nat 139:S4–S35

Schneider H, Sofianidou TS, Kyriakopoulou-Sklavounou P (1988) Calling behavior and calls of *Rana dalmatina* (Anura, Ranidae) in Greece. Zool Physiol Tiere 92(2):231–243

Schull J (1979) A conditioned opponent theory of Pavlovian conditioning and habituation. Psychol Learn Motiv 13:57–90

Schwartz JJ, Simmons AM (1990) Encoding of a spectrally complex natural call in the bullfrog's auditory nerve. J Comp Physiol A 166(4):489–499

Searcy WA (1989) Pseudoreplication, external validity and the design of playback experiments. Anim Behav 38:715–717

Shen JX, Feng AS, Xu ZM, Yu ZL, Arch VS, Yu XJ, Narins PM (2008) Ultrasonic frogs show hyperacute phonotaxis to female courtship calls. Nature 453(7197):914–916

Shen JX, Xu ZM, Yu ZL, Wang S, Zheng DZ, Fan SC (2011) Ultrasonic frogs show extraordinary sex differences in auditory frequency sensitivity. Nat Commun 2:342

Sherman P, Reeve H, Pfennig D (1997) Recognition systems. In: Krebs JR, Davies NB (eds) Behavioural ecology, 4th edn. Blackwell Science, Oxford, pp. 69–96

Shettleworth SJ (2010) Cognition, evolution, and behavior, 2nd edn. Oxford University Press, Oxford

Siddle DA (1991) Orienting, habituation, and resource allocation: an associative analysis. Psychophysiology 28(3):245–259

Simmons AM (2013) "To ear is human, to frogive is divine": Bob Capranica's legacy to auditory neuroethology. J Comp Physiol A 199(3):169–182

Simmons AM, Ferragamo M (1993) Periodicity extraction in the anuran auditory nerve 1. "pitch-shift" effects. J Comp Physiol A 172(1):57–69

Simmons AM, Schwartz JJ, Ferragamo M (1992) Auditory nerve representation of a complex communication sound in background noise. J Acoust Soc Am 91(5):2831–2844

Simmons AM, Reese G, Ferragamo M (1993) Periodicity extraction in the anuran auditory nerve. II phase and temporal fine-structure. J Acoust Soc Am 93(6):3374–3389

Starks PT (2004) Recognition systems: from components to conservation. Ann Zool Fenn 41:689–690

Stoddard PK (1996) Vocal recognition of neighbors by territorial passerines. In: Kroodsma DE, Miller EH (eds) Ecology and evolution of acoustic communication in birds. Cornell University Press, Ithaca, pp. 356–374

Summers K, Symula R, Clough M, Cronin T (1999) Visual mate choice in poison frogs. Proc R Soc Lond B Biol Sci 266(1434):2141–2145

Suthers RA, Narins PM, Lin WY, Schnitzler HU, Denzinger A, Xu CH, Feng AS (2006) Voices of the dead: complex nonlinear vocal signals from the larynx of an ultrasonic frog. J Exp Biol 209 (24):4984–4993

Temeles EJ (1994) The role of neighbors in territorial systems: when are they dear enemies? Anim Behav 47(2):339–350

Thompson RF, Spencer WA (1966) Habituation: a model phenomenon for the study of neuronal substrates of behavior. Psychol Rev 73(1):16–43

Tibbetts EA, Dale J (2007) Individual recognition: it is good to be different. Trends Ecol Evol 22 (10):529–537

Tomsic D, de Astrada MB, Sztarker J, Maldonado H (2009) Behavioral and neuronal attributes of short- and long-term habituation in the crab *Chasmagnathus*. Neurobiol Learn Mem 92 (2):176–182

Tumulty J, Morales V, Summers K (2014) The biparental care hypothesis for the evolution of monogamy: experimental evidence in an amphibian. Behav Ecol 25(2):262–270

Wagner AR (1976) Priming in STM: an information-processing mechanism for self-generated or retrieval-generated depression in performance. In: Tighe TJ, Leaton RN (eds) Habituation: perspectives from child development, animal behavior, and neurophysiology. Lawrence Erlbaum Associates, Hillsdale, pp. 95–128

Wagner AR (1979) Habituation and memory. In: Dickenson A, Boakes RA (eds) Mechanisms of learning and motivation: a memorial volume for Jerzy Konorski. Lawrence Erlbaum Associates, Hillsdale, pp. 53–82

Wagner AR (1981) SOP: a model of automatic memory processing in animal behavior. In: Spear NE, Miller RR (eds) Information processing in animals: memory mechanisms. Lawrence Erlbaum Associates, Hillsdale, pp. 5–47

Wells KD (1977) The social behaviour of anuran amphibians. Anim Behav 25:666–693

Wells KD (2007) The ecology and behavior of amphibians. University of Chicago Press, Chicago

Wells KD, Schwartz JJ (2007) The behavioral ecology of anuran communication. In: Narins PM, Feng AS, Fay RR, Popper AN (eds) Hearing and sound communication in amphibians, Springer handbook of auditory research, vol 28. Springer, New York, pp. 44–86

Whitlow JW, Wagner AR (1984) Memory and habituation. In: Peeke HVS, Petrinovich L (eds) Habituation, sensitization, and behavior. Academic, New York, pp. 103–153

Wiewandt TA (1969) Vocalization, aggressive behavior, and territoriality in the bullfrog, *Rana catesbeiana*. Copeia 1969:276–285

Wiley RH (2013) Specificity and multiplicity in the recognition of individuals: implications for the evolution of social behaviour. Biol Rev 88(1):179–195

Wiley RH (2015) Noise matters: the evolution of communication. Harvard University Press, Cambridge

Wiley RH, Richards DG (1978) Physical constraints on acoustic communication in the atmosphere: implications for the evolution of animal vocalizations. Behav Ecol Sociobiol 3 (1):69–94

Wiley RH, Wiley MS (1977) Recognition of neighbors' duets by stripe-backed wrens *Campylorhynchus nuchalis*. Behaviour 62:10–34

Wilson EO (1975) Sociobiology: the new synthesis. Harvard University Press, Cambridge

Zimmermann E (1990) Behavioral signals and reproduction modes in the neotropical frog family Dendrobatidae. Fortschr Zool 38:61–73

Chapter 8
Referents and Semantics in Animal Vocalizations

Marta B. Manser

Abstract Animal communication is based on signals that provide information to receivers regarding specific aspects of the environment and individual traits of the signaler. Many animals produce acoustically different call types depending on the different behaviors or general contexts they experience. The acoustic structure within a call type typically varies and conveys socially relevant information specific to individual identity, sex, age, social rank, relatedness, or group membership. Both specific referents to the context and referents to individual and group signatures enable receivers to extract diverse information and to incorporate it into their decisions at different levels of complexity in social interactions. From the production side, it is difficult to prove what cognitive mechanisms underlie the emission of specific call types, but recent empirical studies support the fact that it cannot be based on simple emotional expressions. More likely, multiple information processes are involved that integrate the individual traits and an animal's perceptions of different referents, the overall context, or other external stimuli, to produce the final acoustic outcome. Research on the cognitive mechanisms that underlie the perception of different types of referents reveals that on the receiver side, information use likely has both innate and learned components. As such, in all cases, a cognitive representation of the eliciting stimuli expressed by the specific call structure is likely learned by receivers based on simple association of the signal's acoustic structure and the context or the individual traits of the signaler. In the case of functionally referential signals, referents to external stimuli seem to play an influential role in affecting the response of receivers, allowing less flexibility to integrate additional information, compared to other, less context-specific calls, due to the urgency of responding. The different referents in a call should generally reflect the social and ecological constraints a species experiences.

M.B. Manser (✉)
Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland
e-mail: marta.manser@ieu.uzh.ch

M.A. Bee, C.T. Miller (eds.), *Psychological Mechanisms in Animal Communication*, Animal Signals and Communication 5,
DOI 10.1007/978-3-319-48690-1_8

8.1 Introduction

Animals coordinate their activities with conspecifics by communicating with each other and thereby produce a variety of signals relevant to their survival and reproduction, and hence their evolutionary fitness. Communication plays a particularly important role within group-living species in allowing individuals to coordinate their daily activities and form social relationships. Vocal communication is common in many birds (Chaps. 2 and 3) and mammals (Chaps. 9, 10 and 11), although olfactory, visual, and tactile signals are also frequently used (Bradbury and Vehrencamp 2011). While it is obvious that any signal produced has some level of reference (Marler et al. 1992), the debate over which animal vocalizations can be regarded as referential signals remains contentious, and a similarly controversial issue persists regarding the semantics of vocalizations, that is, the meaning of the signal to the receiver (e.g., Stegman 2013; Scarantino and Clay 2015; Wheeler and Fischer 2012, 2015). Referents and semantics in animal vocalizations are the topics of this chapter.

Both of these issues—referents and semantics—were initially introduced into the study of animal communication through research on alarm calling by vervet monkeys, *Chlorocebus pygerythrus* (formerly *Cercopithecus aethiops*). This work described and quantified the production of, and responses to, alarm calls that were specific to different types of predators (Seyfarth et al. 1980; Marler et al. 1992). By confirming Struhsaker's (1967) observations on alarm call production, and by using playback experiments to isolate the context-independent information provided by signals, Seyfarth et al. (1980) provided the first evidence of referential communication in animals (Manser 2013). They not only confirmed that vervet monkeys produce distinct alarm calls for leopards, eagles, snakes, and baboons. They also quantified the distinctive behavioral responses elicited by each of these predator types. For example, vervets ran up into a tree when a leopard appeared. When an eagle appeared, individuals on the ground ran into a bush or tree, whereas individuals already in a tree moved down from its top to the center of the tree out of harm's way. When a snake was encountered, the animals all stood up bipedally and looked around on the ground. In field playback tests, hearing the distinctive alarm calls evoked by each type of predator, broadcast in the absence of the predator itself, elicited the same behavioral responses that were appropriate for the type of predator that had originally elicited the call used as a stimulus. That is, when the animals heard an alarm call originally evoked by a leopard, for example, they responded as if a leopard were actually present. Within the different call types, some variation in call structure (e.g., call length, call interval, or amplitude) appeared to be arousal-related or to express individual traits. Although, for example, an increase in alarm call length increased responsiveness in some cases, these acoustic properties did not affect the qualitative distinctions among responses to the predator-specific call types.

Based on this seminal work on vervet monkeys, "functionally referential" signals have been defined as signals that refer to external objects and events and

Table 8.1 Different types of referents in animal vocalizations and their characteristics with examples

Referent	Call associated with	Context specificity	Examples
Individual traits and social categories (Sect. 8.2.1)	Group membership, age, sex, dominance rank, quality, individual identity, etc.	Varies with breadth of social category and whether or not there are clear individual differences	Any of the calls having variation associated with social category, group membership, or individual status (e.g., "signatures")
Behavioral context (Sect. 8.2.2)	Current behavior of signaler or their motivational or emotional states	Ranges from broad contexts to specific behaviors	Contact calls while foraging
External objects and events (Sect. 8.2.3)	Presence of predators or food	Highly context specific to external stimuli	Alarm calls, food calls

convey a specific semantic meaning to receivers (Macedonia and Evans 1993). In this chapter, I seek to develop a somewhat broader framework for considering referential signals that unifies work on several related topics in the field of animal communication. In addition to including the typical definition of referential signals as referring to external objects and events, this framework, outlined in Sect. 8.2, also recognizes two additional types of referents (Table 8.1). First, it recognizes that vocal signals commonly carry information that refers to the phenotypic traits or social status of individuals, as well as to their membership in various social categories and their social relationships. In Sect. 8.2.1, I discuss studies of recognition and discrimination of individuals and other social categories as well as recognition of third-party social relationships. Second, this framework recognizes that signals can also contain references to the behavioral context and ongoing expressions of specific behaviors. I discuss these issues in Sect. 8.2.2. In Sect. 8.2.3 I return to a discussion of alarm signaling as an example of signals for which external objects and events are the referents. The development of this framework is then followed in Sect. 8.3 by a general discussion of semanticity in animal communication. Section 8.4 briefly considers the psychological mechanisms potentially involved in producing and receiving the three different types of referential signals. Throughout the chapter, I will distinguish between the production and the perception side of signaling behavior and address the question of whether similar cognitive mechanisms underlie communication involving the three broad types of referents in animal vocalizations (Table 8.1). To facilitate the discussion, I will focus on the work we have been doing over the last two decades on meerkats (*Suricata suricatta*) (Fig. 8.1a) and, to a lesser extent, the banded mongoose (*Mungos mungo*) (Fig. 8.1b). In parallel, I integrate results from this work, where possible, with research on other non-primate mammal and bird species and compare it with primates. Before turning to the types of references in communication signals

(a) meerkats (*Suricata suricatta*)

(b) banded mongooses (*Mungos mungo*)

Fig. 8.1 Two mammalian study systems for investigating referents and semantics in animal vocalizations. (**a**) A study group of meerkats. The dominant male wears a radio transmitter that allows researchers to locate the group at any time (photo courtesy Tim Clutton-Brock). (**b**) A study group of banded mongooses (photo courtesy Feargus Cooney)

in Sect. 8.2, I wish to provide a brief background on some of the important issues at hand and on the main study animal to be discussed.

8.1.1 Overview of the Issues

Signals evolve if they bring an advantage to the signaler as well as to the receiver. The signaler influences the behavior or physiology of the receiver to its own advantage (Maynard-Smith and Harper 2003), and the receiver typically responds in a way that is of advantage to itself. In cooperative situations, the interests of signalers and receivers overlap, while in conflict situations, signalers and receivers have different motivations in the production of the signal and how to respond, and benefits to the communicative partners may differ substantially. In general, the mechanisms underlying the production of signals by the signaler differ from the mechanisms that are involved in the perception of the signal and subsequent generation of a response (Seyfarth and Cheney 2003). Vocal signals had long been regarded as motivational or emotional expressions of animals (Darwin 1872; Morton 1977). Following the study describing alarm calls specific to different predator types in vervet monkeys (Struhsaker 1967), however, researchers realized that specific external stimuli can elicit highly context-specific calls. Subsequently, the discussion emerged as to whether these calls refer to external events or objects (Seyfarth et al. 1980; Marler et al. 1992; Macedonia and Evans 1993) and whether receivers have a cognitive representation of the eliciting stimuli that can be evoked by only hearing the vocalizations (Zuberbühler et al. 1999). However, recent discussions about the cognitive mechanisms underlying vocal perception question whether these highly context-specific calls induced by an external event or object should be regarded as different than calls associated with a specific behavior or the individual traits of the signaler (Wheeler and Fischer 2012).

Vocal signals for most mammal species have been described as innate and hardwired, leaving little room for adjustment and flexibility in different social and ecological contexts. This is particularly true for the production side with regard to the signal's overall acoustic structure, but is less true for the usage and comprehension of calls (Smith 1965, 1981; Chaps. 9 and 10). So far only a few species have been documented to adjust their calls as adults and, for example, to conform as a social group to a vocal signature shared in common (humpback whales, *Megaptera novaeangliae*, Payne and Payne 1985; chimpanzees, *Pan troglodytes*, Crockford et al. 2004) or to imitate conspecifics (bottlenose dolphins, *Tursiops truncatus*, King and Janik 2013) or heterospecifics (harbor seals, *Phoca vitulina*, Ralls et al. 1985; Asian elephants, *Elephas maximus*, Stoeger et al. 2012). More frequently, mammal and bird species have been reported to exhibit high flexibility in tailoring signal usage to their social environment (Seyfarth and Cheney 2010). Similarly, from the perception side, we observe considerable variability in the likelihood that receivers respond to the same call types and also in the strength of their responses.

Regarding underlying cognitive mechanisms, these patterns suggest differences between call production, call usage, and call comprehension. In many species, call production seems primarily genetically determined and triggered by specific external or internal factors, without much control and flexibility on the part of the signaler. Whereas in call usage and call comprehension, the signaler and receiver show much more flexibility and capacity to adjust their signals to their social and ecological environment (see Chap. 10). Therefore, the key questions we are interested in addressing are what causes variation in the acoustic structure of a signaler's calls, and what aspects of this variation do receivers perceive as meaningful in terms of changing their behavior? The follow-up questions are then, what are the underlying cognitive mechanisms on the production side, in terms of call usage, and on the perception side, in terms of generating a response, and in what ways do they differ? We are, in particular, interested in asking these questions separately as they pertain to the three different types of references in animal vocalizations mentioned above (Table 8.1) and outlined in more detail in Sect. 8.2.

8.1.2 Meerkats

Meerkats, a cooperatively living mongoose species of the family Herpestidae, live in despotic societies with the dominant pair monopolizing reproduction (Fig. 8.1a; Clutton-Brock et al. 1998). They occupy the open, dry habitat of a semidesert, with scarce food availability and high predation pressure (Clutton-Brock et al. 2001). The group typically consists of between 3 and 50 individuals, with several subordinate adult and subadult individuals, juveniles, and pups (Clutton-Brock et al. 2006). Each group defends a territory of about 2–5 km^2 (Manser and Bell 2004) by marking along their territory boundaries, and other important locations, using feces and anal gland secretion (Jordan et al. 2007). The dominant female can produce up to 4 litters a year with 1–7 pups per litter (Clutton-Brock et al. 2001). Dominant females contribute up to 80 % of the pups in the population, and the dominant male sires about 80 % of the dominant female's pups (Griffin et al. 2003; Spong et al. 2008; Nielsen et al. 2012). Subordinate individuals typically forego their own reproduction and help raise the dominant pair's offspring (Clutton-Brock et al. 1998), although under some circumstances, they are able to reproduce and even raise their own offspring. A group can be highly stable over many years and have the same dominant pair, but in other groups, dominant pairs, or at least one of the dominant individuals, may change frequently, bringing instability to the group, often with the result of decreased reproduction (unpublished data, long-term Kalahari Meerkat Project).

The dominant pair exhibits distinctive behaviors to assert their dominance position, such as by frequent anal markings at obvious locations within their home range, as well as by regularly marking other group members. Subordinates periodically show clear submissive behaviors by initiating grooming and emitting specific vocalizations (Kutsukake and Clutton-Brock 2006). A recent analysis focusing on social interactions among subordinate females has shown that there is also subtle competition and that a social hierarchy exists, whereby age and condition seem to be the determining factors in what relative ranking a specific individual assumes (Thavarajah et al. 2014).

Meerkats have evolved a rich repertoire of signals in several different modalities that are used to coordinate their cooperative behavior and cohesive group movement, as well as in forming social relationships. This includes a sophisticated vocal system (Fig. 8.2; Manser et al. 2014), in addition to olfactory (Jordan et al. 2007) and visual signals. They produce at least 30 different call types, and, particularly in the context of predator avoidance, calls are associated with high variation in their acoustic structures (Manser et al. 2014). Due to their cohesive foraging, and the fact that a single individual must trade off being vigilant for predators against digging

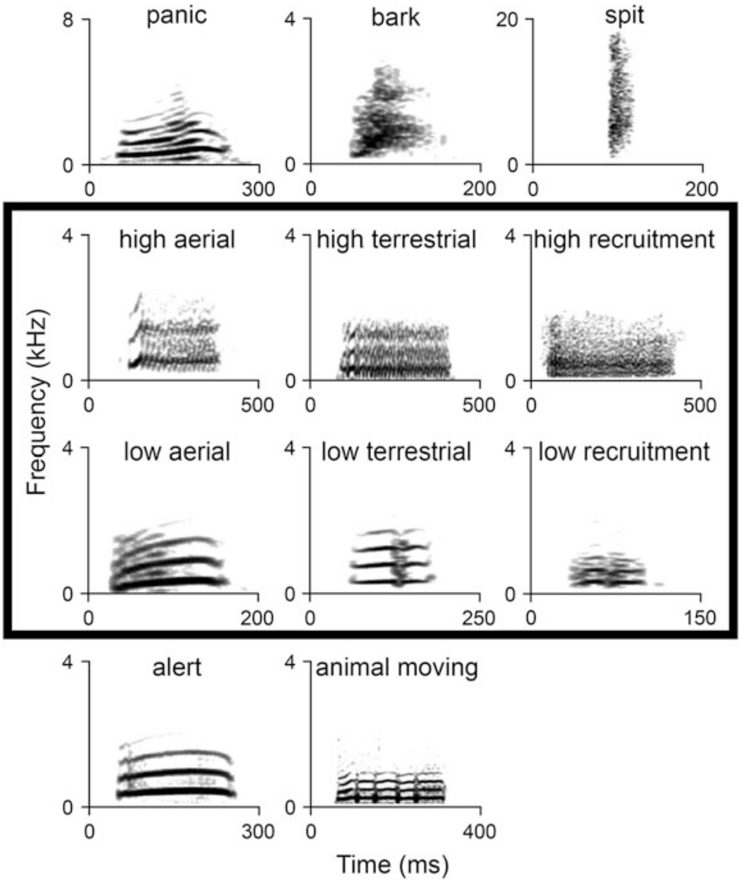


Fig. 8.2 Spectrograms of the different alarm calls of meerkats (*Suricata suricatta*). *Top row* shows the highest-urgency calls; *bottom row*, the lowest-urgency calls, which are not predator-type specific. *Middle rows* (framed) show the predator-specific calls (aerial, terrestrial, recruitment) in relation to the urgency (low, high) of the situation (Reproduced from Manser 2009)

for prey in the sand, they also coordinate their movement while foraging using several different call types. Finally, calls also play an important role in both affiliative and agonistic social interactions. Pup vocalizations differ from the adult vocal repertoire (White 2001). In particular, several different types of begging calls are produced to elicit provisioning by the older group members (Manser and Avey 2000; Kunc et al. 2007). After individuals reach approximately 6 months of age, the full range of adult vocalizations is in place (Hollén et al. 2008).

Olfactory, visual, and tactile cues and signals also play critical roles in meerkat communication, though to date they have been less thoroughly investigated. Olfactory cues or signals include the deposition of feces, urine, and scent marks from different glands, but mainly anal glands (Jordan et al. 2007). These signals are used to maintain the dominance hierarchy within a group, signal territory boundaries, and potentially advertise the reproductive state of females within the group. Visual signals include different body postures, but also glares that engage others visually, or changes in tail positions (M. B. Manser, personal observations) that, in particular, occur during social interactions.

8.2 Types of Referential Vocal Signals

Animal communication has been defined as functionally referential if a signaler produces a signal that is associated with a specific object or event in the external environment and allows the receiver to extract that information and use it in generating an accordingly specific response (Marler et al. 1992; Macedonia and Evans 1993). Calls related to the behavioral state of the signaler were regarded as the expression of the motivation or emotion of the caller (Darwin 1872; Morton 1977; for review Manser 2009). Recent discussions have questioned whether a distinction of functionally referential calls in comparison to other less context-specific call types or calls referring to the signaler's behavioral state is justified, based on the cognitive mechanisms involved in perceiving and responding to signals (Wheeler and Fischer 2012). The basic argument is that any call type, whether highly context specific to the external environment of the signaler or referring to the signaler's behavioral state, conveys some referential information to the receiver. In this section, I describe an expanded framework for referential signals that considers how acoustic variation in calls relates to different types of references in vocal communication in terms of conveying specific information to receivers (Table 8.1).

8.2.1 Reference to Individual Traits and Social Categories

Phenotypic variation of traits related to the individuality or social category of the signaler allows social recognition at different levels (Wiley 2013; Chap. 7). This variation, thus, conveys a type of reference to the receiver by differentiating animals according to their group membership, kin, rank, sex, age, or quality, with the most fine-grained categorization occurring at the individual level, or even within an individual (e.g., due to hormonal changes related to reproductive stage or health condition). Testing whether individual differences in vocalizations are meaningful to the receiver has challenged researchers. One particular recurring hurdle is whether receivers just *discriminate* among individuals or categories of individuals or whether they *recognize* the individual and have some cognitive representation of it (Proops et al. 2009).

Categorizing conspecifics is important for any animal but in particular for animal species that live in groups with repeated encounters and individualized interactions. In such social settings, it becomes advantageous to distinguish among individuals based on differences in relationships (Wittig et al. 2007) or contribution to cooperative tasks (Krams et al. 2008). In almost all mammal species, animals produce individually distinctive vocalizations and olfactory signals due to differences in their morphology or physiology, irrespective of whether there has been past selection for individuality (Sheehan et al. 2014). In this section, I will evaluate the different types of social recognition in general, and in meerkats in particular, to identify the underlying cognitive mechanisms.

8.2.1.1 Individual Discrimination and Recognition

Despite the fact that individuality in most vocalizations has been described early on, such as Beecher (1982, 1991) emphasizing when individuality is likely to be evolutionarily advantageous, it has taken a long time to collect evidence that individually specific signals are used by receivers. In the past, researchers typically recorded calls of several individuals and quantified the differences in vocalizations. Most of the time, this was done from one sample of recordings, and we often do not know how stable individual traits in calls are in relation to the ontogeny of an animal. Some call structures vary also within an individual, for example, due to hormonal influences related to dominance rank, condition, or health state, such as cortisol or testosterone having a direct influence on pitch or call rate and length.

The strength of selection for individuality depends on interactions among group members, the group composition, and in particular also on the group size. In species where group members aggregate but do not show individualized interactions, recognition is less important than in groups where such individualized interactions have evolved. In such species, being individually recognized is more difficult in larger groups (Pollard and Blumstein 2011). Therefore, in social animals, for which recognizing different individuals brings benefits, increased group size may select for increased individuality. Evidence in support of this hypothesis exists in analyses of contact calls of bats (Wilkinson 2003) and alarm calls in sciurid rodents (Pollard and Blumstein 2011). It remains to be investigated whether similar relationships exist in the different social mongoose species, which vary in maximal group size, with dwarf mongoose groups having up to 30 members, those of meerkats ranging up to 50 individuals, and those of the banded mongoose extending up to 70 individuals (Manser et al. 2014). Also, the question arises as to whether these effects of group size may also apply within a species, with the calls in larger groups becoming more individually distinct with smaller within-individual variation and larger among-individual variation. Furthermore, some types of vocalizations, either due to their acoustic structure or their function and the context they are related to, may be more individually distinct than others (Rendall et al. 2009).

Methods to identify how animals categorize their social environment differ depending on the level of recognition required. Individual discrimination is based on a simple mechanism, where receivers detect acoustic differences between the calls of different individuals. In contrast, individual recognition requires receivers to have a cognitive representation about which specific individual is associated with different call structures (Seyfarth and Cheney 2015). Individual or categorical discrimination can be tested by simple habituation-discrimination experiments (Johnston and Bullock 2001; Chap. 7), whereas the question of individual or categorical recognition is better investigated by violations of expectation experiments (Proops et al. 2009). In habituation-discrimination experiments, a subject is presented with a set of different stimuli of the same category until it no longer shows a continued response (Johnston and Bullock 2001). Once habituation has occurred, a stimulus of a different category is presented. If this new stimulus causes a response similar to that at the very beginning, we have evidence that the animals distinguish between the two stimulus types. If the subject's response does not recover, it becomes difficult to interpret the result (see discussions of this issue in Chaps. 6 and 7). For example, did the subject fail to respond because the difference between stimuli was smaller than the corresponding just-noticeable difference (JND)? Did the difference exceed the subject's JND but fall below its just-meaningful difference (JMD; Nelson and Marler 1990); that is, were the differences perceptually discriminable but behaviorally irrelevant? Or was the experimental setup somehow flawed or simply not adequately realistic or representative? To exclude failures of discrimination due to failures of experimental design, it is imperative to simultaneously run appropriate control conditions (Hare and Atkins 2001; Schibler and Manser 2007; Karp et al. 2014).

One of the most convincing studies on individual recognition of vocalizations was conducted with domestic horses (*Equus caballus*) by employing a cross-modal expectancy violation experiment (Proops et al. 2009). In expectancy violation experiments, a cue or signal is presented that goes together with previous information about an event or detected object (congruent situation). Then a test is performed to determine whether the animal responds in a surprised manner if the presented cue or signal does not correspond to the original information (incongruent situation). In the study by Proops et al. (2009), subjects were first presented with a specific horse from their group who was then led behind a visual barrier. This generated a visual expectancy regarding which horse was blocked from view. Using playbacks of horse contact calls from the direction of the barrier, which were either congruent or incongruent with the horse hidden behind the barrier, this expectancy was either confirmed or violated. If the vocalization and the horse the test subject was exposed to were congruent, the response was minimal. In the situation where the horse vocalization being broadcast was from a different individual, the test subject showed surprised behavior. These results have since been confirmed in other animals using similar setups testing cross-modality between vocal and visual cues (large-billed crows, *Corvus macrorhynchos*, Kondo et al. 2012; rhesus monkeys, *Macaca mulatta*, Adachi and Hampton 2011) and also between vocal and olfactory cues (ring-tailed lemurs, *Lemur catta*, Kulahci et al. 2015).

The maintenance of the social structure displayed by meerkats seems to require individual recognition and not just individual discrimination. This recognition could occur within the visual, olfactory, or vocal modalities. With experiments we have shown that they vocally distinguish among individuals and that this has to be at the recognition level rather than just discriminating among calls of different individuals (Townsend et al. 2012a; Reber et al. 2013). With an expectancy violation experiment in the spatial arrangement of foraging meerkats, we first showed that meerkats discriminate between calls of different individuals (Townsend et al. 2012a). In a follow-up experiment, we showed that subordinate meerkats also distinguish between any other subordinate female in the group and the dominant female (Reber et al. 2013). Here, we controlled that it was not just a dominance expression in the acoustic structure of the calls but rather the recognition of the dominant females. Based on these two experiments, we concluded that meerkats recognize individuals from their calls, at least their "close calls," by recognizing the acoustic variation and associating it with specific group members.

Individual discrimination or recognition is not always used by receivers even if individual differences in signals exist. In meerkats we have shown in several different contexts that calls vary individually, including alarm calls (Schibler and Manser 2007), sentinel calls (Manser 1998), and close calls (Townsend et al. 2010). However, the receivers seem not to use this information in all circumstances. When testing individual discrimination of alarm calls with a habituation-discrimination experiment, meerkats showed the same response, whether it was an individual that had been made unreliable or another reliable group member (Schibler and Manser 2007). This indicated that the identity of the animal producing the alarm call was either not perceived or was not important in a receiver's decision to respond. Such a lack of discrimination may be understandable in a high-risk situation such as alarm calling. Although in other species, such as marmots (*Marmota marmota*), receivers take the identity of the signaler into account in similar high risk situations (Blumstein and Daniel 2004). In meerkats, unreliable signalers may be less frequent due to the high predation risk, such that they cannot afford to cheat (Schibler and Manser 2007).

Another situation showing the flexibility of receivers is the context dependence of meerkat responses to close calls. Close calls are the most frequently emitted calls while foraging, with a large variation in call rate, but also individual variation in call structure, as well as more extensive group differences (Townsend et al. 2010). The response of receivers to the dominant female's close calls described in the above experiment only becomes obvious when there is a conflict between the listener and the dominant female, but not during stable, relaxed periods (Reber et al. 2013). Similar results have been reported in chacma baboons (*Papio ursinus*) where a female responded to another female's threat grunt only if she had recently been threatened by that female. If she had recently groomed with that same female, she ignored the call (Engh et al. 2006). These experiments clearly show that receivers discriminate individual differences in vocalizations and that they are also able to represent the signaler in ways that go beyond the memory of their individually distinct signals. However, if there is no motivation to show a response, we are not able to distinguish the causes of no response. It then becomes difficult to judge whether individual differences are not important or the experimental setup was not appropriate to the species' natural behavior, as may have been the case when we tested the same setup in the banded mongoose, and no variation in response was detected (Jansen et al. 2013).

8.2.1.2 Social Category Discrimination and Recognition

Social recognition is not only about individual recognition but also occurs at other levels, such as group, kin, rank, sex, age, or the advertisement of quality. Several species have evolved group signatures in their vocalizations and scents. In theory, such group signatures facilitate the immediate identification of the group a signaler belongs to and, as such, can be used to help advertise or defend home ranges or group resources. Group-specific calls may be due to genetic differences, as shown for Gunnison's prairie dogs (*Cynomys gunnisoni*, Travis et al. 1997) and squirrel monkeys (*Saimiri sciureus*, Lieblich et al. 1980), or due to vocal learning, as shown in humpback whales (*M. novaeangliae*, Payne and Payne 1985), yellow-naped amazons (*Amazona auropalliata*, Wright and Wilkinson 2001), and chimpanzees (*P. troglodytes*, Crockford et al. 2004).

In meerkats we have shown that individuals in different groups vary in vocal and olfactory signals individually, but also appear to have a group signature (Townsend et al. 2010; Wadewitz 2010). However, while we found a strong response to olfactory cues of foreign meerkats (Wadewitz 2010), the animals did not show much interest in the calls of foreign individuals (Townsend et al. 2010; Reber et al. 2013). The same result was found for the banded mongoose, in which subjects respond particularly strongly to the olfactory cues of neighbors and less strongly to those of strangers (Müller and Manser 2007), but when playing contact calls of neighbors, they do not respond at all (Müller 2007).

The group signature in meerkat vocalizations is unlikely to be under strong selection given that it is not used by receivers to distinguish between their own group and foreign groups (Townsend et al. 2010). Olfactory signals, however, may be much more important in these contexts. This is not surprising if we consider that vocalizations are used mainly during within-group communication and not in between-group interactions or in advertisements to potential mates. The one context where it may be beneficial to adjust vocalizations to other group members is for male immigrants, so as not to be too different from others in the newly joined group and, thus, to facilitate being recognized as a group member. However, as long as having a distinct individual signature is not disadvantageous, there may be no selection pressure to change it, and conformity in meerkats, due to social or ecological factors, may not be expected. Ongoing work is currently investigating how immigrant adult males adjust their individual signature in calls given in their new social group. This may differ in regard to olfactory signals, as they are used to communicate within and between groups, and this is true for meerkats (Jordan et al. 2007; Wadewitz 2010) and banded mongooses (Jordan et al. 2010).

Kin also represent an important social category for which advertisement and recognition can be beneficial. This is particularly true in mate choice contexts, in which inbreeding avoidance is paramount, or in cooperative breeding contexts, in which investing help in closely related individuals is key to increasing indirect fitness benefits. For bell miners (*Manorina melanophrys*), a colonial honeyeater from Australia, the calls of more closely related individuals are more similar in comparison with unrelated individuals, though helping effort correlated not with genetic relatedness but with acoustic similarity (McDonald and Wright 2011). While evidence for kin recognition of vocalizations in mate choice is missing, olfactory signals play an important role in mate preference (Le Claire et al. 2013). Several different mechanisms may facilitate kin recognition, including familiarity (i.e., growing up in close spatial proximity with ample opportunities for social interactions) or phenotype matching, in which animals assess their own phenotype or that of the phenotype of their familiar kin and compare it to the encountered unfamiliar individuals (Lacy and Sherman 1983; Le Vin et al. 2010).

Meerkats live in social groups with several overlapping generations and in which mainly the females are philopatric and the males disperse typically at the age of 1–3 years (Clutton-Brock et al. 1998; Young and Clutton-Brock 2006). Females disperse after they have been forced out of their natal group by the dominant female and then sometimes meet up with males to found a new group. Within groups, meerkat shows clear inbreeding avoidance, in which the dominant female does not breed with her sons, the dominant male does not breed with his daughters, and full-siblings or half-siblings do not breed with each other (Nielsen et al. 2012). The question was how dispersing females and males avoid inbreeding, as they might encounter full-siblings or half-siblings they had never encountered before and, thus, could not rely on familiarity as a mechanism for inbreeding avoidance. The most obvious signals to use are olfactory, as vocal signals are not frequently given in such encounters, and visual cues, although there seem to be some similarities among group members, may not be as reliable. In a study testing kin recognition in meerkats via olfactory signals, in particular anal gland secretions, females invested more time inspecting the scent of related than unrelated unfamiliar individuals, suggesting that they use a phenotype-matching mechanism to discriminate kin from non-kin (Le Claire et al. 2013).

The recognition of social rank may be based on simple mechanisms where each individual in the group divides its social companions into two groups: those ranked above itself and those ranked below. Slightly more complicated mechanisms may involve an individual understanding of the rank relations among every other individual in the group, which has been shown in chacma baboons (*Papio ursinus*, Bergman et al. 2003) and pinyon jays (*Gymnorhinus cyanocephalus*, Paz-y-Miño et al. 2004). Here, rank order recognition may involve transitive inference, meaning that subjects recognize if A dominates B and B dominates C, then A dominates C. In meerkats we have the clear distinct role of the dominant pair, where within the same sex, they are dominant to everybody else. Among the subordinates, a rank order also exists and appears to be related to age given that, within litter/age cohorts, same-sex siblings establish their relative rankings by showing dominance assertion to each other (Thavarajah et al. 2014). Existing data suggest meerkats accept older individuals as higher ranked, but compete for ranks within their cohort. If the dominant position were freed up, the most dominant individual in the next cohort has an advantage because it can quickly assume dominance without long, costly competition for the position against other same-sex individuals in the group.

Recognition of other social categories, such as sex or quality, seems not to play an important role in the vocal behaviors of meerkats. This contrasts with some other mammal species, in which, for example, we find clear differences between females and males in the structure of alarm calls (e.g., vervet monkeys, Seyfarth and Cheney 1990; green monkeys, *Chlorocebus sabaeus*, Price and Fischer 2013), suggesting that the function of alarm calls may be different for the two sexes. That we do not find a sex difference in meerkats might be related to the fact that we have little evidence that quality advertisement plays a role in recruiting mates. However, quality advertisement could be important in the direct competition within the different sexes, when competition is most obvious. This may be the case during dominance changes within each sex, when individuals compete very obviously with each other (Clutton-Brock et al. 2006; Kutsukake and Clutton-Brock 2006), but also among subordinates during relaxed periods, when there seems to be a continuous assertion of dominance, in particular among litter mates (Thavarajah et al. 2014). We are currently addressing these questions to see whether the most frequently emitted close calls or any other vocalization types change in individuals during the transition from subordinate to dominant. Such a change would not be surprising, as females experience a secondary growth period during this stage (Russell et al. 2004). The question would still remain as to whether such changes are at all meaningful to receivers.

8.2.1.3 Recognition of Third-Party Relationships

Recognition of relationships among other group members, that is, third-party relationships, is considered to be a particularly complex cognitive challenge (Seyfarth and Cheney 2015). Recognizing third-party relationships requires more complex cognitive mechanisms than the ability to individually recognize conspecifics. It requires observing and appropriately interpreting third-party interactions and integrating this information into one's future behavioral decisions. Even in small groups, and in particular in larger groups, where a large number of dyadic or even triadic interactions are possible, a substantial memory of individualized interactions is required. Several studies of chacma baboons (Cheney and Seyfarth 1999) and also of spotted hyenas (*Crocuta crocuta*, Engh et al. 2005) have provided evidence for the recognition of third-party relationships. In meerkats, it is currently unclear whether convincing evidence for the recognition of relationships among other group members exists, except for the relationship between dominant female and male as the dominant pair. The problem may be partly that we are not able to detect such relationships, as there are no obvious contexts in which specific individuals seem to support each other more or have a stronger relationship than with others.

8.2.2 References to Behavioral Contexts

Many animals vocally express their current behavior with specific vocalizations emitted on a regular or ongoing basis while they are engaged in the behavior. Such calls can be related to broad contexts, such as foraging as a cohesive group. In some species, variation in call structure maps directly onto specific events such as searching and feeding events (Jansen et al. 2012) or travelling between food patches (Boinski 1991). Some acoustically very similar calls, which we may categorize as the same call type, are used in several different behavioral contexts to which a signaler is exposed or according to the behavior it performs. For example, in chacma baboons, acoustically similar calls are used as contact barks while foraging and as alarm barks to warn others of predators (Fischer et al. 2001). Many other behaviors are much more confined to a specific context but are also accompanied by vocalizations, such as aggression in competitive interactions over resources, as well as various affiliative behaviors, including grooming.

Such calls have been suggested to express the motivation and also the emotional state of the signaler (Darwin 1872; Morton 1977). For example, an imminent attack might be signaled with harsh aggressive calls, or submission might be signaled when social partners express fear in order to avoid likely aggression. Based on these observations, Morton (1977) put forward some empirical evidence that signal design follows so-called “motivational-structural rules” in that a signaler’s emotions or motivations are clearly represented by specific acoustic features. Several studies in birds and mammals provided additional empirical support for this, but other studies did not find this pattern (Manser 2009). On the perception side, this means that receivers need additional information on context to distinguish different situations when callers emit acoustically similar calls purely on an emotional basis.

In meerkats, the majority of behavioral contexts are accompanied by a specific vocalization, which may explain the large repertoire of distinct call types they have evolved (Manser 1998; Manser et al. 2014). The main contexts in which vocalizations are emitted involve coordinating group cohesion (e.g., while foraging and moving in their home range), avoiding predation (e.g., predator warnings, mobbing, recruiting others to inspect predator cues, or coordinating sentinel activities), and interacting socially with conspecifics (e.g., affiliative or aggressive behaviors). Some of the same call types are used in several different contexts. For example, the same call types are given during sentinel activities and when individuals are standing in the sun to warm themselves, as well as during grooming. Other distinct call types are limited to one context only. Lead calls, for example, are only given when an individual wants to move and the group to follow; a submission call is only emitted when the animal submits to a more dominant individual. Currently we do not fully understand why some call types occur across contexts, in particular when they are associated with broad categories of potentially very different emotions. One explanation could be that meerkats categorize the calls differently than we do based on our available quantitative analytical methods.

8.2.3 References to External Objects and Events

Following the early studies of alarm calling in vervet monkeys, several studies have now demonstrated that other mammal and bird species produce highly object-specific and context-specific vocalizations that refer to features of the external environment, primarily in the context of predator alarm calls and food calls (reviewed in Townsend and Manser 2013; Gill and Bierema 2013). In meerkats, many of the identified call types are associated with a specific behavior of the signaler (Manser 1998; Manser et al. 2014). In addition, like vervet monkeys, meerkats emit predator-type specific calls that are restricted specifically to the approach of predators and that induce stereotyped escape behaviors with some flexibility along an urgency continuum (Manser 2001; Manser et al. 2001, 2002). They also have a call that is elicited by the detection of other animals in their environment, but only when moving, independent of whether the detected animals are dangerous or harmless birds or mammals, foreign conspecifics, or predators (Manser 2009). Besides their predator-type specific terrestrial and aerial alarm calls, meerkats produce other alarm calls that have a more general alert function, and receivers need additional contextual information to perceive the details of the situation (Manser 2001, 2009) (Fig. 8.2).

The acoustic structures of predator-type specific calls in meerkats vary depending on how close the threat is and what risk it poses (Manser 2001; Manser et al. 2001, 2002). This means the calls convey information to the receiver in regard to multiple aspects of the situation. Firstly, the call refers to the predator type, although it is not clear whether this information is about the predator identity or the spatial area and direction from which it approaches (e.g., aerial raptors approaching from the sky versus terrestrial predators approaching on the ground). Secondly, the acoustic structure changes along the same dimension within the predator-specific calls conveying information on the distance and the risk the approaching predator poses (Manser 2001). These multidimensional aspects of variation related to different external factors make it difficult to distinguish whether calls refer to the external event or are the expression of the emotional state of the signaler, as the whole discussion on predator-type specific calls versus urgency-level-based alarm calls has shown (e.g., Furrer and Manser 2009).

To be considered a functionally referential call, a high degree of perception specificity must exist on the part of receivers, as demonstrated by their appropriate response to the call in the absence of the stimulus that elicits the call from signalers (Marler et al. 1992). In meerkats, playbacks of different predator-type specific calls (see Fig. 8.2) caused them to respond in qualitatively different ways (Manser et al. 2001). In response to an aerial alarm call, they would run to the next shelter, typically a bolthole. In response to a terrestrial alarm call, they would look around and gather together to then either move to a sleeping burrow system for shelter or, if no predator was to be seen, resume foraging. In response to a recruitment call, receivers would move towards the caller or loudspeaker. When testing the different predator-type specific calls that differed in their urgency level, which is clearly expressed in changes in acoustic structure (Fig. 8.2), receivers showed quantitative changes in the intensity of their responses, but responses remained qualitatively the same within a predator type.

Emitting an alarm call is not only affected by the external stimuli eliciting the call but also by the social environment of the caller in meerkats (Townsend et al. 2012b) and also in other species (Papworth et al. 2008). Such findings suggest that, even though the emission of alarm calls may generally be a rather stereotyped, genetically determined behavior, there is some flexibility on the production side, and different information processes are involved in the decision of whether to give a predator-type specific call. Likewise, on the perception side, not all receivers respond in the same way. Instead, we find variation among individuals in terms of whether they respond, how fast and strong they respond, and how fast they resume foraging after hearing an alarm call. In particular, it appears that individuals that have invested time into digging for a prey item in the sand are more reluctant to run immediately for shelter, and if they do run, they are the first to resume foraging again and return to their digging spot (Amsler 2008). These observations support the idea that behavioral responses to a call with a specific acoustic structure are not simple reflexes. Instead, receivers appear to incorporate additional information into their decisions about responding.

8.3 Semanticity in Animal Communication

The debate over the potential referential and semantic nature of signals must be considered from both the production and perception sides (Macedonia and Evans 1993; Seyfarth and Cheney 2003). The highly context-specific alarm calls of vervet monkeys were initially termed *functionally* referential (Marler et al. 1992) to emphasize that the underlying cognitive mechanisms driving production and perception may well be different from those underlying the referential use of words in human language. For many years, people have questioned whether animal communication involves information transfer or instead involves manipulation of receivers by signalers (Dawkins and Krebs 1978; Stegmann 2013). Owings and Morton (1998) built up their "assessment-management" approach to vocal communication based on Morton's (1977) motivational structural rules. However, there is increasing evidence that the production of call types differing in usage and structure cannot simply be explained as merely the expression of the signaler's motivational and emotional state, but rather involves multiple information processes (Manser 2009; Crockford et al. 2012, Watson et al. 2015).

The perception side of functionally referential vocalizations has also been reconsidered in recent years. Owren and Rendall (1997) put forward the "affect conditioning" or "affect induction" model, where animals do not respond in a specific way because of the information being transferred in a signal, but due to the effect of the signal's specific acoustic structure on the receiver's low-level perceptual, attentional, and motivational processes. However, there is also plenty of evidence that the affect model has its limitations, for example, when it comes to explaining the flexible responses of receivers to the same acoustic structure of vocalizations (Seyfarth et al. 2010), and that semantics and referential information likely play important roles in generating receiver responses.

Arguments over the semantic properties of functionally referential signals are based on their purported effects in evoking cognitive representations of the eliciting objects or events in the mind of the receiver. The clearly varied responses to the various alarm call types in vervet monkeys raises questions about whether a simple association between a specific call type and an external stimulus might allow receivers to respond appropriately (perceptual semanticity) or whether an evoked cognitive representation of the signal-eliciting stimulus induces the response in the receiver (conceptual semanticity) (Zuberbühler et al. 1999). There is evidence from prime-probe experiments on Diana monkeys (*Cercopithecus diana*) that habituation was transferable across semantically similar calls but not across acoustically similar ones (Zuberbühler et al. 1999). Such findings suggest the animals were not responding to the acoustic features alone; instead, their responses were mediated by the similarity of the meaning of the presented stimuli. Evans and Evans (2007) supported this idea with experiments on chickens with regard to their food calls being representational signals. Nevertheless, the common associations in both of these examples could be formed via associative learning and may not necessarily involve more complex cognitive representations (Adams and Beighley 2013).

If we accept that animals are able to recognize individuals from their vocalizations in a cross-modal, representational way (Proops et al. 2009), it could be argued that functionally referential food and predator vocalizations are processed in a similar way. Specifically, responses to such vocalizations may be mediated by a receiver's cognitive representation of the object eliciting the call type (Zuberbühler et al. 1999; Hurford 2007). In all cases of different referent types (Table 8.1), animals appear capable of learning to make these types of associations. This view is supported by experiments with golden-mantled ground squirrels (*Spermophilus lateralis*, Shriner 1999) and, more recently, by experiments on superb fairy wrens (*Malurus cyaneus*), in which predators were associated with artificial sounds (Magrath et al. 2015). The receivers showed clear escape behavior after few exposures, whereas they did not show a response to a new call brought in without association to a predator.

Another recent suggestion discourages distinguishing in a *categorical* way between functionally referential calls and behavior-related calls in favor of considering all animal vocalizations as existing along a *continuum* from low to high context specificity (Wheeler and Fischer 2012). These arguments are based on the cognitive processes involved from the perception side, with one of the main assumptions being that, to the receiver of the calls, there is no inherent difference between those associated with external referents or internal features of the signaler. This may be true for the perception side, but a critical question remains on the production side: what underlying cognitive processes cause the emission of the specific call types in the first place? Currently, we seem unable to convincingly identify the relevant processes involved at the level of what specific aspects of external stimuli drive the production of specific call types, much less at a neurobiological level. Neither do we fully understand the cognitive processes involved in the production of calls relating to specific behaviors of a caller. Clearly much work remains before we will fully understand the semanticity of signals from the perspectives of both signalers and receivers.

8.4 Mechanisms for Producing and Receiving Referential Signals

In this section, I briefly address issues related to the underlying mechanisms involved in producing and receiving referential signals as they pertain to the three types of referential signals discussed in Sect. 8.2 and summarized in Table 8.1.

From the production side, considering the sources and patterns of variation in signals is key to gaining insight into their production mechanisms. For example, the variation in signals that is correlated with differences in individual traits may often be due to anatomical and physiological differences related to sound production mechanisms (see Chap. 7). That is, the vocal expression of individual traits may be more or less fixed because production mechanisms limit the signaler's influence on the signal's individually distinctive acoustic structure, perhaps depending on the signal type (e.g., noisy versus tonal) or also its function (Rendall et al. 2009). Consequently, many of the signals animals produce will necessarily include some reference to individual traits or membership in various social categories. This form of referential signaling, therefore, likely occurs independently of whether the signals are also used to convey information about specific behavioral contexts or external objects and events.

It is more difficult to identify mechanisms driving signalers to produce calls that reference specific behavioral contexts or external objects and events. In terms of signaling-specific behavioral contexts, variation in signals may arise due to variation in the signaler's own behavioral or motivational state, which may show large variability depending on the signaler's immediate environment. For example, to emphasize a situation as becoming threatening in fights over resources, signalers may first start with low levels of aggression and then change into high levels of aggression. In turn, these state-level changes may be reflected in an increase in the modulation, harshness, and amplitude of signals (Manser et al. 2014). In many instances, however, adjustments in signal production appear to reflect the signaler's ability to attend to differences in its social environments, often in ways that depend on its own situation and the situation of one or more intended receivers. Indeed, there is increasing empirical evidence that, instead of being a simple expression of emotion or motivation, the production and usage of signals is more flexible than previously described. Signalers tailor the production of their calls to the social environment, including both whether and what categories of conspecifics are around (Townsend et al. 2012b; Gyger et al. 1987), whether or not receivers attend to the signal (Wich and de Vries 2006), and whether or not conspecifics are in danger (Papworth et al. 2008). For example, there is evidence from predator-type specific calls (Townsend et al. 2012b) and food-type specific calls (Gros-Louis 2004) that signalers do, in fact, adjust their signaling behavior in relation to their social environment. This necessarily means that signalers may often possess cognitive mechanisms for processing information about their social environment and making decisions that give them flexibility to change their vocal production based not only on the behavior and responses of receivers but also, at least in some species, on the knowledge they perceive the receiver to possess (Crockford et al. 2012; see also Chap. 9), implying a qualitatively different level of social awareness (but see Seyfarth and Cheney 2015). A signaler that tries to influence the receiver to produce a specific response may be at a significant advantage if it is able to anticipate the action of the receiver and change its own vocal behavior accordingly.

On the perception side, receiver mechanisms for processing information in referential signals and generating appropriate behavioral responses appear to depend on the type of referent (Table 8.1) to which receivers benefit most from attending. As noted previously, referents to behavioral contexts and also to external objects and events all seem to also include referents to individual traits. Yet, the attention of receivers to calls referring to external objects and events seems primarily biased to this dimension and thus less to individual traits. This bias, in turn, accordingly guides the behavior of the receiver to be more stereotyped and less dependent on the signaler's identity. This bias and its behavioral effects are likely due to signals referring to external objects and events that are commonly associated with high arousal (Meise et al. 2011) and a high urgency to generate the most appropriate response. Predator-type specific calls related to danger, for example, induce fear. Food-type specific calls related to high rewards induce positive excitement or the desire to move to the location of a food source. Regarding signals referring to the signaler's behavioral context, there is seldom the need to respond as immediately as to functionally referential calls, and the cognitive integration of the reference on individual traits, as well as other context-related information, can easily become more important. This would support the suggestion by Wheeler and Fischer (2012) that in less context-specific signals, animals need to acquire additional information to decide how best to respond, and this may be cognitively more demanding. However, it may be overestimated that in highly context-specific calls, other information sources may be fully filtered out, as variation in the responses of receivers to functionally referential calls suggests flexibility in several species (Seyfarth et al. 1980; Manser et al. 2001; Price and Fischer 2013).

8.5 Summary and Future Directions

Producing and receiving referential vocal signals, at least in social mammals and birds, should be considered in a broader framework that recognizes three different types of referents (Table 8.1): those related to individual traits and social categories, those related to behavioral context, and those related to external objects and events. In this chapter, I have illustrated the benefits of adopting such a framework using examples largely drawn from our work on vocal communication and social behavior in meerkats. The meerkat vocal system seems to be characterized by an exceptionally broad range of predator-specific (external referent) and behavior-specific (pertaining to the signaler) vocalizations. This may be explained by the fact that meerkats use their vocalizations mainly to coordinate group cohesion and antipredator behaviors, both of which are keys to survival in the open habitat they live in with scarce food availability and high predation pressure. Although meerkat social interactions are also accompanied by specific vocalizations, dyadic interactions in meerkats seem to be more important in organizing social relationships than attending to third-party relationships, as shown for more socially structured groups in primates or hyenas. Olfactory and visual signals and cues also play important roles in meerkat social interactions by providing additional, multimodal information about the characteristics of an individual, thereby freeing them from sole reliance on mainly vocal signals (see Chap. 5). The predator-type specific calls produced by signalers elicit distinct and obvious escape responses from receivers that are qualitatively consistent and appropriate to the type of predator. Nevertheless, variation in the strength of a response depends on the behavior of the receiver at the moment the signal is perceived.

To understand the function and meaning of vocalizations, future efforts will need to identify both the causes of the acoustic variation observed in the production of the signals and the consequences of that variation in terms of the responses of receivers. Being aware that we have different references in vocalizations, with each based on different underlying cognitive mechanisms at least on the production side, though potentially less so on the perception side, may help to make clear predictions for what variation we can expect in specific situations. For example, the demands of social recognition will differ depending on the social system or social structure of an organism and will select more or less for the expression of individual traits or the coordination of different behavioral tasks. Ecological constraints, such as those related to predation risk and foraging, may differentially favor the evolution of signals referring to external object or events.

While over the recent years a lot of progress has been made toward identifying in detail acoustic variation as it relates to specific referents, we still do not know much about the specific underlying information processes related to the production or the perception of acoustic variation. For example, we have only limited knowledge about the conditions that enhance selection for individual variation or uniformity and what acoustic parameters are typically expressing them under what conditions. We still do not fully understand what acoustic structures elicit specific behaviors, or why in some species the same signal type is used in several behavioral contexts, while in other species, several different signal types seem to elicit the same behaviors. In terms of the so-called functionally referential signals that refer to external objects and events, we generally have not adequately identified whether the calls really refer to spatial area versus predator type, whether the signals are given as a command, or whether they merely express the caller's emotional state.

Likewise, on the perception side, we still lack a complete understanding about what information processes underlie responses to specific variation in the acoustic features of signals. Although, evidence is slowly emerging that some animals do seem to have cognitive representations for individuality and predator types, empirical studies are very rare, and it is difficult to draw firm or broad conclusions about what cognitive mechanisms for information processing and integration are involved.

The biological world is very seldom divided into distinct black and white categories, even though it may be much easier for us to comprehend and quantify discrete rather than continuous structures and patterns. It is, therefore, not surprising that we as observers tend to focus on the most obvious, or on the very novel and exciting aspects of a system, and suppress the additional "noisy" effects that are difficult to place in the conceptual frameworks we develop to make sense of the world. For example, the first papers on the functionally referential alarm calls of vervet monkey focused on the predator specificity of signals and the qualitatively different responses of receivers, but largely disregarded the variation in signals and responses in relation to other factors (Seyfarth et al. 1980). Recent studies on the same species highlight this variation and suggest the calls are not as predator-type specific as originally described (Price et al. 2015).

While we gain from advancing single concepts as the main explanation, it becomes increasingly clear that animals are exposed to a vast variety of inputs that subsequently function to stimulate signal production and, on the receiver side, function to make the best decision in their situation to respond to specific vocalizations. One aspect I have learned from our detailed work on communication and cognition in meerkats is that we are still far from understanding how these animals categorize their world and in what way this influences their communication system. To understand animals' decisions and the underlying cognitive mechanisms, it is critical that we first identify in detail what input from the social and ecological environment is relevant and influences the production of signals and responses to them.

Acknowledgments I thank the editors for the invitation to write this chapter and in particular Mark Bee for his editing efforts. Also many thanks to Simon Townsend and two referees on their helpful comments on earlier drafts of the chapter. The University of Zurich has funded this project and also the long-term research on the different mongoose species in addition to the Swiss National Science Foundation.

References

Adachi I, Hampton R (2011) Rhesus monkeys see who they hear: spontaneous cross-modal memory for familiar conspecifics. PLoS One 6(8):e23345

Adams F, Beighley SM (2013) Information, meaning and animal communication. In: Stegmann UE (ed) Animal communication theory: information and influence. Cambridge University Press, Cambridge, pp. 399–420

Amsler V (2008) How urgency levels in alarm calls influence the foragers response in meerkats (*Suricata suricatta*). MSc thesis, University of Zurich, Zurich

Beecher MD (1982) Signature systems and kin recognition. Am Zool 22(3):477–490

Beecher MD (1991) Successes and failures of parent-offspring recognition systems in animals. In: Hepper PG (ed) Kin recognition. Cambridge University Press, Cambridge, pp. 94–124

Bergman T, Beehner JC, Cheney DL, Seyfarth RM (2003) Hierarchical classification by rank and kinship in baboons. Science 302:1234–1236

Blumstein DT, Daniel JC (2004) Yellow-bellied marmots discriminate between the alarm calls of individuals and are more responsive to calls from juveniles. Anim Behav 68(6):1257–1265

Boinski S (1991) The coordination of spatial position: a field study of the vocal behaviour of adult female squirrel monkeys. Anim Behav 41(1):89–102

Bradbury JW, Vehrencamp SL (2011) Principles of animal communication, 2nd edn. Sinauer Associates, Sunderland, MA

Cheney DL, Seyfarth RM (1999) Recognition of other individuals' social relationships by female baboons. Anim Behav 58:67–75

Clutton-Brock TH, Brotherton PNM, Smith R, McIlrath GM, Kansky R, Gaynor D, O'Rian MJ, Skinner JD (1998) Infanticide and expulsion of females in a cooperative mammal. Proc R Soc Lond B 265(1412):2291–2295

Clutton-Brock TH, Brotherton PNM, Russell AF, O'Riain MJ, Gaynor D, Kansky R, Griffin A, Manser M, Sharpe L, McIlrath GM, Small T, Moss A, Monfort S (2001) Cooperation, control, and concession in meerkat groups. Science 291(5503):478–481

Clutton-Brock T, Hodge S, Spong G, Russell A, Jordan N, Bennett N, Sharpe L, Manser M (2006) Intrasexual competition and sexual selection in cooperative mammals. Nature 444 (7122):1065–1068

Crockford C, Herbinger I, Vigilant L, Boesch C (2004) Wild chimpanzees produce group-specific calls: a case for vocal learning? Ethology 110(3):221–243

Crockford C, Wittig RM, Mundry R, Zuberbühler K (2012) Wild chimpanzees inform ignorant group members of danger. Curr Biol 22:142–146

Darwin CH (1872) The expression of the emotions in man and animals. Murray J, Publ, London

Dawkins R, Krebs JR (1978) Animal signals: information or manipulation. In: Krebs JR, Davies NB (eds) Behavioural ecology: an evolutionary approach. Blackwell Scientific Press, Oxford, pp. 282–309

Engh AL, Siebert ER, Greenberg DA, Holekamp K (2005) Patterns of alliance formation and postconflict aggression indicate spotted hyenas recognize third-party relationships. Anim Behav 69:209–217

Engh AL, Hoffmeier RR, Cheney DL, Seyfarth RM (2006) Who, me? Can baboons infer the target of a vocalization? Anim Behav 71:381–387

Evans CS, Evans L (2007) Representational signalling in birds. Biol Lett 3(1):8–11

Fischer J, Metz M, Cheney DL, Seyfarth RM (2001) Baboon responses to graded bark variants. Anim Behav 61(5):925–931

Furrer R, Manser MB (2009) The evolution of urgency-based and functionally referential alarm calls in ground-dwelling species. Am Nat 173:400–410

Gill SA, Bierema AMK (2013) On the meaning of alarm calls: a review of functional reference in avian alarm calling. Ethology 119:449–461

Griffin AS, Pemberton JM, Brotherton PNM, McIlrath G, Gaynor D, Kansky R, O'Riain J, Clutton-Brock TH (2003) A genetic analysis of breeding success in the cooperative meerkat (*Suricata suricatta*). Behav Ecol 14(4):472–480

Gros-Louis J (2004) The function of food-associated calls in white-faced capuchin monkeys, *Cebus capucinus*, from the perspective of the signaller. Anim Behav 67:431–440

Gyger M, Marler P, Pickert R (1987) Semantics of an avian alarm call system: the male domestic fowl (*Gallus domesticus*). Behaviour 102:15–40

Hare JF, Atkins BA (2001) The squirrel that cried wolf: reliability detection by juvenile Richardson's ground squirrels (*Spermophilus richardsonii*). Behav Ecol Sociobiol 51:108–112

Hollén L, Clutton-Brock TH, Manser MB (2008) Factors affecting the development of alarm-call production, usage and responses in meerkats (*Suricata suricatta*). Behav Ecol Sociobiol 62:821–829

Hurford JR (2007) The origins of meaning: language in the light of evolution. Oxford University Press, Oxford

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

Jansen DAWAM, Cant MA, Manser MB (2013) Testing for vocal individual discrimination in adult banded mongooses. J Zool 291:171–177

Johnston RE, Bullock TA (2001) Individual recognition by use of odours in golden hamsters: the nature of individual representation. Anim Behav 61:545–557

Jordan NR, Cherry MI, Manser MB (2007) Latrine distribution and patterns of use by wild meerkats: implications for territory and mate defence. Anim Behav 73(4):613–622

Jordan NR, Mwanguhya F, Kyabulima S, Rüedi P, Cant MA (2010) Scent marking within and between groups of wild banded mongooses. J Zool 280(1):72–83

Karp D, Manser MB, Wiley EM, Townsend SW (2014) Nonlinearities in meerkat alarm calls prevent receivers from habituating. Ethology 120(2):189–196

King SL, Janik VM (2013) Bottlenose dolphins can use learned vocal labels to address each other. Proc Natl Acad Sci 110(32):13216–13221

Kondo N, Izawa E, Watanabe S (2012) Crows cross-modally recognize group members but not non-group members. Proc R Soc Lond B 279:1937–1942

Krams I, Krama T, Igaune K, Mänd R (2008) Experimental evidence for reciprocal altruism in the pied flycatcher. Behav Ecol 62:599–605

Kulahci IG, Drea CM, Rubenstein DI, Ghazanfar AA (2015) Individual recognition through olfactory–auditory matching in lemurs. Proc R Soc Lond B 281(1784):20140071

Kunc HP, Madden JR, Manser MB (2007) Begging signals in a mobile feeding system: the evolution of different call types. Am Nat 170(4):617–624

Kutsukake N, Clutton-Brock TH (2006) Aggression and submission reflect reproductive conflict between females in cooperatively breeding meerkats, *Suricata suricatta*. Behav Ecol Sociobiol 59(4):541–548

Lacy R, Sherman PW (1983) Kin recognition by phenotype matching. Am Nat 121:489–512

Le Claire S, Nielsen JF, Thavarajah NK, Manser M, Clutton-Brock TH (2013) Odour-based kin discrimination in the cooperatively breeding meerkat. Biol Lett 9(1):20121054

Le Vin AL, Mable BK, Arnold KE (2010) Kin recognition via phenotype matching in a cooperatively breeding cichlid, *Neolamprologus pulcher*. Anim Behav 79:1109–1114

Lieblich AK, Symmes D, Newman JD, Shapiro M (1980) Development of the isolation peep in laboratory-bred squirrel monkeys. Anim Behav 28(1):1–9

Macedonia JM, Evans CS (1993) Essay on contemporary issues in ethology: variation among mammalian alarm call systems and the problem of meaning in animal signals. Ethology 93:177–197

Magrath RD, Haff TM, McLachlan JR, Igic B (2015) Wild birds learn to eavesdrop on heterospecific alarm calls. Curr Biol 25:2047–2050

Manser, M.B. (1998) The evolution of auditory communication in suricates *Suricata suricatta*. PhD thesis, Cambridge University, Cambridge

Manser MB (2001) The acoustic structure of suricate alarm calls varies with predator type and the level of urgency. Proc R Soc Lond B 268:2315–2324

Manser MB (2009) What do functionally alarm calls refer to? In: Dukas R, Ratcliffe J (eds) Cognitive ecology II. The University of Chicago Press, Chicago, IL and London, pp. 229–248

Manser MB (2013) Semantic communication in vervet monkeys and other animals. Anim Behav 86:491–496

Manser MB, Avey G (2000) The effect of pup vocalisations on food allocation in a cooperative mammal, the meerkat (*Suricata suricatta*). Behav Ecol Sociobiol 48(6):429–437

Manser MB, Bell MB (2004) Spatial representation of shelter locations in meerkats, *Suricata suricatta*. Anim Behav 68(1):151–157

Manser MB, Bell MB, Fletcher LB (2001) The information that receivers extract from alarm calls in suricates. Proc R Soc Lond B 268:2485–2491

Manser MB, Seyfarth RM, Cheney DL (2002) Suricate alarm calls signal predator class and urgency. Trends Cogn Sci 6:55–57

Manser MB, Graw B, Hollén LI, Bousquet CA, Furrer RD, le Roux A (2014) Vocal complexity in meerkats and other mongoose species. Adv Stud Behav 46:281–310

Marler P, Evans CS, Hauser MD (1992) Animal signals: reference, motivation, or both? In: Papousek H, Jürgens U, Papousek M (eds) Nonverbal vocal communication: comparative and developmental approaches. Cambridge University Press, Cambridge, pp. 66–86

Maynard-Smith JM, Harper D (2003) Animal signals. University Press, Oxford, New York, NY

McDonald PG, Wright J (2011) Bell miner provisioning calls are more similar among relatives and are used by helpers at the nest to bias their effort towards kin. Proc R Soc Lond B 278 (1723):3403–3411

Meise K, Keller C, Cowlishaw G, Fischer J (2011) Sources of acoustic variation: implications for production specificity and call categorization in chacma baboon (*Papio ursinus*) grunts. J Acoust Soc Am 129:1631–1641

Morton ES (1977) On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. Am Nat 111:855–869

Müller, C.A. (2007) Environmental knowledge in the banded mongoose (*Mungos mungo*). PhD thesis, University of Zurich, Zurich

Müller CA, Manser MB (2007) 'Nasty neighbours' rather than 'dear enemies' in a social carnivore. Proc R Soc Lond B 274(1612):959–965

Nelson DA, Marler P (1990) The perception of birdsong and an ecological concept of signal space. In: Berkley MA, Stebbins WC (eds) Comparative perception: volume II. Wiley, New York, NY, pp. 443–478

Nielsen JF, English S, Goodall-Copestake WP, Wang J, Walling CA, Bateman AW, Flower TP, Sutcliffe RL, Samson J, Thavarajah NK, Kruuk LE, Clutton-Brock TH, Pemberton JM (2012) Inbreeding and inbreeding depression of early life traits in a cooperative mammal. Mol Ecol 21 (11):2788–2804

Owings DH, Morton ES (1998) Animal vocal communication: a new approach. Cambridge University Press, Cambridge

Owren MJ, Rendall D (1997) An affect-conditioning model of nonhuman primate vocal signaling. Perspect Ethol 12:299–346

Papworth S, Böse AS, Barker J, Schel AM, Zuberbühler K (2008) Male blue monkeys alarm call in response to danger experienced by others. Biol Lett 4:472–475

Payne K, Payne R (1985) Large-scale changes over 19 years in song of humpback whales in Bermuda. Z Tierpsychol 68:89–114

Paz-y-Miño CG, Bond AB, Kamil AC, Balda RP (2004) Pinyon jays use transitive inference to predict social dominance. Nature 430:778e782

Pollard KA, Blumstein DT (2011) Social group size predicts the evolution of individuality. Curr Biol 21:413–417

Price T, Fischer J (2013) Meaning attribution in the West African green monkey: influence of call type and context. Anim Cogn 17:277–286

Price T, Hammerschmidt K, Wadewitz P, Cheney DL, Seyfarth RM, Fischer J (2015) Vervets revisited: a quantitative analysis of alarm call structure and context specificity. Sci Rep 5:13220

Proops L, McComb K, Reby D (2009) Cross-modal individual recognition in domestic horses (*Equus caballus*). Proc Natl Acad Sci USA 106(3):947–951

Ralls K, Fiorelli P, Gish S (1985) Vocalizations and vocal mimicry in captive harbor seals: *Phoca vitulina*. Can J Zool 63:1050–1056

Reber SA, Townsend SW, Manser MB (2013) Social monitoring via close calls in meerkats. Proc R Soc Lond B 280:1–7

Rendall D, Notman H, Owren MJ (2009) Asymmetries in the individual distinctiveness and maternal recognition of infant contact calls and distress screams in baboons. J Acoust Soc Am 125(3):1792–1805

Russell AF, Carlson AA, McIlrath GM, Jordan NR, Clutton-Brock T (2004) Adaptive size modification by dominant female meerkats. Evolution 58(7):1600–1607

Scarantino A, Clay Z (2015) Contextually variable signals can be functionally referential. Anim Behav 100:e1–e8

Schibler F, Manser MB (2007) The irrelevance of individual discrimination in meerkat alarm calls. Anim Behav 74:1259–1268

Seyfarth R, Cheney D (1990) The assessment by vervet monkeys of their own and another species' alarm calls. Anim Behav 40(4):754–764

Seyfarth RM, Cheney DL (2003) Signalers and receivers in animal communication. Annu Rev Psychol 54:145–173

Seyfarth RM, Cheney DL (2010) Production, usage, and comprehension in animal vocalizations. Brain Lang 115(1):92–100

Seyfarth RM, Cheney DL (2015) Social cognition. Anim Behav 103:191–202

Seyfarth RM, Cheney DL, Marler P (1980) Vervet monkey alarm calls: semantic communication in a free-ranging primate. Anim Behav 28:1070–1094

Seyfarth RM, Cheney DL, Bergman T, Fischer J, Zuberbühler K, Hammerschmidt K (2010) The central importance of information in studies of animal communication. Anim Behav 80:3–8

Sheehan MJ, Straub MA, Tibbetts EA (2014) How does individual recognition evolve? Comparing responses to identity information in *Polistes* species with and without individual recognition. Ethology 120(2):169–179

Shriner WM (1999) Antipredator responses to a previously neutral sound by free-living adult golden-mantled ground squirrels, *Spermophilus lateralis* (Sciuridae). Ethology 105:747–757

Smith WJ (1965) Message, meaning and context in ethology. Am Nat 99:405–409

Smith WJ (1981) Referents of animal communication. Anim Behav 29:1273–1274

Spong GF, Hodge SJ, Young AJ, Clutton-Brock TH (2008) Factors affecting the reproductive success of dominant male meerkats. Mol Ecol 17(9):2287–2299

Stegmann U (2013) Animal communication theory: information and influence. Cambridge University Press, Cambridge

Stoeger AS, Mietchen D, Oh S, de Silva S, Herbst CT, Kwon S, Fitch WT (2012) An Asian elephant imitates human speech. Curr Biol 22(22):2144–2148

Struhsaker TT (1967) Auditory communication among vervet monkeys (*Cercopithecus aethiops*). In: Altmann SA (ed) Social communication among primates. University of Chicago Press, Chicago, IL, pp. 281–324

Thavarajah NK, Fenkes M, Clutton-Brock TH (2014) The determinants of dominance relationships among subordinate females in the cooperatively meerkat. Behaviour 151:89–102

Townsend SW, Manser MB (2013) Functionally referential communication in mammals: the past, present and the future. Ethology 119:1–11

Townsend SW, Hollén LI, Manser MB (2010) Meerkat close calls encode group significant signatures, but fail to discriminate. Anim Behav 80:133–138

Townsend SW, Allen C, Manser MB (2012a) A simple test of vocal individual recognition in wild meerkats. Biol Lett 8:179–182

Townsend SW, Rasmussen M, Clutton-Brock TH, Manser MB (2012b) Flexible alarm calling in meerkats: the role of the social environment and predation urgency. Behav Ecol 23:1360–1364

Travis SE, Slobodchikoff C, Keim P (1997) DNA fingerprinting reveals low genetic diversity in Gunnison's prairie dog (*Cynomys gunnisoni*). J Mammal:725–732

Wadewitz P (2010) Olfactory discrimination of predators and conspecifics in meerkats (*Suricata suricatta*). MSc thesis, University of Zurich, Zurich

Watson SK, Townsend SW, Schel AM, Wilke C, Wallace EK, Cheng L, West V, Slocombe KE (2015) Vocal learning in the functionally referential food grunts of chimpanzees. Curr Biol 25 (4):495–499

Wheeler BC, Fischer J (2012) Functionally referential signals: a promising paradigm whose time has passed. Evol Anthropol Issues News Rev 21:195–205

Wheeler BC, Fischer J (2015) The blurred boundaries of functional reference: a response to Scarantino & Clay. Anim Behav 100:e9–e13

White SM (2001) Juvenile development and conflicts of interest in meerkats. PhD thesis, University of Cambridge, Cambridge

Wich SA, de Vries H (2006) Male monkeys remember which group members have given alarm calls. Proc R Soc Lond B 273:735–740

Wiley RH (2013) Specificity and multiplicity in the recognition of individuals: implications for the evolution of social behaviour. Biol Rev 88(1):179–195

Wilkinson GS (2003) Social and vocal complexity in bats. In: de Waal FBM, Tyack PL (eds) Animal social complexity: intelligence, culture, and individualized societies. Harvard University Press, Cambridge, MA, pp. 322–341

Wittig RM, Crockford C, Seyfarth RM, Cheney DL (2007) Vocal alliances in chacma baboons, *Papio hamadryas ursinus*. Behav Ecol Sociobiol 61:899–909

Wright TF, Wilkinson GS (2001) Population genetic structure and vocal dialects in an amazon parrot. Proc R Soc Lond B 268:609–616

Young AJ, Clutton-Brock T (2006) Infanticide by subordinates influences reproductive sharing in cooperatively breeding meerkats. Biol Lett 2(3):385–387

Zuberbühler K, Cheney DL, Seyfarth RM (1999) Conceptual semantics in a nonhuman primate. J Comp Psychol 113:33–42

Chapter 9
Social Concepts and Communication in Nonhuman Primates

Klaus Zuberbühler

Abstract A major question in science concerns how humans evolved their capacity for language. One approach to answering this question is to take the entire faculty of language as the evolutionarily relevant unit and, by applying natural selection theory, look for evidence of variability, heritability, and adaptive function. Another approach is to conceive of the language faculty as a conglomerate of components with independent evolutionary histories and identify phylogenetic continuities and discontinuities among various components. Here, I follow the second approach by focusing on comparative research on primate social cognition and communication. Primates have unusually large brains, and the mainstream view is that large brains are an evolved response to prevail in socially complex worlds. Complexity arises from intergroup and intragroup relations governed by reproductive interests, which are mediated by dominance, kinship, and friendship. This chapter reviews field studies that have addressed these topics, focusing on how primates communicate in the wild to ensure their reproductive interests and in relation to evolutionarily important external events, such as food discoveries or predator encounters. The conclusion from this research is that primates can attribute basic mental states to others, such as intentions or perceptions, but the data are less compelling for more complex mental states, such as beliefs or knowledge. I conclude by proposing a research agenda to investigate in more detail the major evolutionary transitions that have paved the way to the emergence of the human language faculty.

K. Zuberbühler (✉)
Institute of Biology, Department of Comparative Cognition, University of Neuchâtel, Neuchâtel, Switzerland

School of Psychology and Neuroscience, University of St. Andrews, St. Andrews, Scotland, UK
e-mail: klaus.zuberbuehler@unine.ch

M.A. Bee, C.T. Miller (eds.), *Psychological Mechanisms in Animal Communication*, Animal Signals and Communication 5,
DOI 10.1007/978-3-319-48690-1_9

9.1 Brain and Intelligence

The purpose of this chapter is to review behavioral evidence, mainly from natural populations, to elucidate how nonhuman primates organize their worlds internally. Results are relevant for a number of key questions in science, particularly those pertaining to how and why humans have evolved their capacity for language. Human language is based on conceptual thought, and it is, therefore, relevant to understand how nonhuman species organize and represent their worlds internally. Two basic assumptions are worth stating here. First, following Darwin (1871/1998), we can assume there are no fundamental differences in the cognitive capacities between humans and what some might call the higher animals. Second, we can assume there is an intrinsic relationship between relative brain size and cognitive complexity. Given these assumptions, a major question in science has been what could have favored the evolution of large brains in primates in general and humans in particular?

Brain tissue is energetically expensive to maintain, so what are selective pressures that have been responsible for its evolution? As little as two million years ago, *Homo habilis*, a direct ancestral form of the Hominidae family, had an estimated brain size only slightly larger than a modern chimpanzee (*Pan troglodytes*; Tobias 1971), illustrating the rapid evolutionary changes that can take place towards bigger brains and higher intelligence. However, brain evolution is not just about changes in size but also in organization and neural density. For example, the relative size of cortical areas differs between mammalian taxa due to specializations for sensory niches (Barton 2007). An interesting illustration is the human cerebellum, which contains several times more neurons than the neocortex (Azevedo et al. 2009) and appears to be the result of an evolutionarily recent expansion (Barton and Venditti 2014). As the cerebellum is responsible for motor control (Wolf et al. 2009), an intriguing hypothesis is that demands in the physical domain have generated an added selection pressure on human cognitive evolution.

Overall, however, the mainstream view still is that brain size and organization are an evolved response to the advanced demands of living in complex social structures (Humphrey 1976). The “social intelligence hypothesis” has widespread support in many scientific disciplines, with generally good evidence for a causal relationship between social complexity and neocortex size, especially among primate species (Barton and Dunbar 1997) and sometimes even within species. For example, in one recent study, it was found that long-tailed macaques (*Macaca fascicularis*) living in large groups had more gray matter and neural connections in some cortical areas than individuals living in small groups (Sallet et al. 2011). At the same time, it needs to be pointed out that there is another group of theories, which posits that brain evolution actually went the opposite way. Presumably, advanced cognitive abilities are advantages to any individual of any species, but only individuals that can provide the necessary energetic resources can afford large brains (Zuberbühler and Janmaat 2010). It is thus possible that nonhuman primates managed to outcompete other species in controlling access to high quality food in their habitats, and so had the extra resources necessary to evolve large brains.

If large and functionally organized brains are a response to the challenges of social life, as predicted by the social intelligence hypothesis, then primates should be able to organize their worlds along social concepts, not unlike humans do. To address this hypothesis, this chapter reviews a range of behavioral studies of primate cognition largely from natural populations (Zuberbühler 2014). Section 9.2 explores the nature of nonhuman primates' social concepts, such as "leader," "friend," or "enemy." In Sect. 9.3, I review work aimed at demonstrating how these social concepts are manifest in the daily use of communication behavior in primates. Section 9.4 examines the importance of learning in terms of how cognitive representations of social concepts are formed, and how individuals learn to communicate about them. In Sect. 9.5, I briefly take up the question of whether primates perceive others as having cognitive states that may differ from their own. Finally, in Sect. 9.6, I outline an agenda for future research that highlights the likely benefits to come from studying social concepts and communication in natural primate populations.

9.2 Social Concepts in Primates

9.2.1 Concepts of Group Identity

Primates live in groups that are defined by individualized membership. Some of the first systematic work on this was carried out with free-ranging vervet monkeys (*Chlorocebus pygerythrus*, formerly *Cercopithecus aethiops*) in East Africa (Seyfarth and Cheney 1982). These studies demonstrated that individuals not only classify others based on kinship and matrilines but also possess a notion of group membership as shown by the fact that monkeys associated the vocalizations of particular individuals with particular groups (Cheney and Seyfarth 1982). The primate concept of group identity most likely also has a spatial component. For example, in vervet monkeys, it has been shown experimentally that individuals mentally track the movements of others because subjects reacted differently to plausible and implausible displacements simulated by sound playbacks (Noe and Laporte 2014).

Similar results are available for free-ranging chimpanzees. If the sudden presence of another chimpanzee was simulated by playing back his pant-hoot vocalizations from a concealed speaker (Herbinger et al. 2009), subjects reacted with a wide range of gestural signals, but there were significant differences depending on whether the calls were from a familiar group male, a neighboring male, or a stranger male. For example, calls of familiar group members mainly triggered pant-hoot replies, whereas neighbor and stranger males elicited screams. Overall, males responded more strongly than females, and male party size had an effect on subjects' vocal responses. Overall, the study demonstrated that chimpanzees identified other individuals based on their group membership and were aware of the potentially lethal consequences associated with encountering strangers (Wilson et al. 2014). In a related study, Wilson et al. (2001) showed that when encountering strangers, individuals carried out a numerical assessment since responses to playbacks of "pant-hoots" by an extra-group male depended on the number of adult males in the recipient party.

Despite well-documented out-group hostility in chimpanzees and other primates, migration between groups occurs regularly, suggesting that migrating individuals face a difficult task when trying to integrate into a new group. In most primate species, migration is by young adult males, who try to replace the resident male or take over a group where the resident male has disappeared. In multi-male groups, immigrating males try to establish themselves alongside the resident males, obtain high rank, and become highly aggressive towards resident females with young offspring (Palombit 1999).

In a few primate species, including chimpanzees, however, migration is by nulliparous females entering another group to start reproducing (Langergraber et al. 2014). Several studies have documented the social difficulties that young immigrant chimpanzee females face during this life-changing event, especially because of the resident females' hostility towards them. In the Sonso community of Budongo Forest, for example, immigrant females have been subject to severe aggression from older, well-established resident females, who are prepared to exert lethal violence towards the newborn infants of immigrant females, suggesting that the immigrants are under strong selection pressure to evolve counterstrategies (Townsend et al. 2007). One consequence is that, during estrous, females produce their copulation calls strategically so as to either conceal or advertise their sexual activity and likelihood of conception. For example, chimpanzee females call significantly more when copulating with high-ranking males, presumably to increase promiscuous matings and to confuse paternity, but they also suppress their calls if high-ranking females are nearby, presumably to conceal their sexual activity (Townsend et al. 2008).

In sum, concepts of group identity are difficult to investigate directly, but the behavioral evidence in the wild suggests individuals pursue various behavioral strategies to avoid the potentially lethal consequences of intraspecies hostility towards strangers. The topic needs more research, but it is clear that the consequences of intergroup competition and the challenges of migration generate a considerable selection pressure, which is likely to have favored advanced cognitive capacities.

9.2.2 Concepts of Dominance

For many primate species, social units are characterized by relatively clear and linear dominance hierarchies. In ground-breaking field experiments, Bergman et al.

(2003) showed that chacma baboons (*Papio ursinus*) recognize that the dominance hierarchy of their group is additionally organized into family groups, suggesting a hierarchically structured social concept of dominance and kinship. In playback experiments, subjects responded more strongly to call sequences mimicking dominance rank reversals between families compared to within families, regardless of rank differences, indicating that they classify others simultaneously according to both individual rank and kinship. This finding has also been replicated for vervet monkeys, but there were interesting sex differences (Borgeaud et al. 2013). While there was evidence that females knew both the male and female hierarchy, the adult males and juveniles seemed to lack such knowledge (Borgeaud et al. 2015).

Dominance is a driving social force in many species, and one interesting problem is how individuals learn their own and others' positions within a group. Transitive inference has been highlighted as one cognitive capacity by which animals can learn to represent group members within a social hierarchy, a cognitive mechanism that appears to be relatively widespread throughout the animal kingdom, with evidence coming from species as diverse as lemurs (Tromp et al. 2015), fish (Grosenick et al. 2007), and corvids (Emery and Clayton 2004).

Another way of exploring the social dominance hierarchy is by proactively challenging other group members. Male bonobos (*Pan paniscus*) show such behavior by producing acoustically distinct "contest hoots," which are directed at specific individuals and regularly combined with gestures and other body signals. Contest hooting, with or without other signals, functions to provoke a social reaction in another group member, usually an agonistic chase, and males are highly selective in whom they target. Opponents are chosen by their rank, and since individuals of equal or higher social rank are preferred, the suggestion is that this vocal behavior functions to explore or assert social status and, in doing so, showcase their social status to other group members (Genty et al. 2014).

9.2.3 Concepts of Social Bonds

In addition to dominance relations, primate groups are characterized by further social structures, notably kin relations and social bonds between unrelated individuals. In nonhuman primates, social bonds are typically biased towards kin, especially in species with large multi-male or multi-female groups, such as macques and baboons, and there is evidence that individuals recognize the kin relations of other group members. For example, using a match-to-sample task, Dasser (1988) showed that long-tailed macaques were able to differentiate "mother–offspring" dyads from other dyads, even if subjects were tested with images of group members at a much younger age.

But in many species there is also evidence that individuals maintain differentiated personal relationships beyond mother–offspring bonds and other kin-related structures (i.e., "friendships"). For example, in Diana monkeys (*Cercopithecus diana*) and Campbell's monkeys (*Cercopithecus campbelli*), individuals preferentially target group members other than their closest relatives to form social bonds (Candiotti et al. 2015). More generally, primate social bonds can occur within and between the sexes, can last for years, and involve cooperative interactions separated in time (Seyfarth and Cheney 2012). Although reproductive pair bonds are common in many species, primates are able to maintain social bonds with individuals other than their reproductive partners. The ability to form social bonds with nonrelatives may thus be one of the most important ways in which primate sociality differs from that found in most other birds and mammals. Dunbar and Shultz (2007) argue that the demands of forming and maintaining social bonds were the critical factor that triggered the evolution of large brains, rather than the more traditional view of living in large, complex societies.

There is a general trend in that for species with male dispersal (most primates), friendships are more likely among females while in species with female dispersal, friendships are more likely among males (Seyfarth and Cheney 2012). Social bonds provide a selective advantage for males in that they may give individuals superior competitive abilities and improve their reproductive success. For females, strong and enduring social bonds result in lower levels of stress, higher infant survival, and longer life expectancy (Seyfarth and Cheney 2012). Recent research in chimpanzees has demonstrated that there is a physiological component to social bonding. Oxytocin levels were higher after grooming with bond partners compared with non-bond partners, regardless of genetic relatedness or sexual interest (Crockford et al. 2013). This finding also has implications on how primates and other social animals maintain cooperative relationships. One explanation has been in terms of some form of mental bookkeeping of cooperative exchanges, but the chimpanzee (Crockford et al. 2013) study puts the emphasis on endocrinological mechanisms that act directly on neural reward and social memory systems.

Another interesting problem is not only how social bonds are formed, but also how other group members view and respond to the bonds of others. In one of the first documented field experiments addressing this problem, Kummer et al. (1974) investigated in a set of enclosure experiments the early phases of bond formation between male and female hamadryas baboons (*Papio hamadryas*). In one experiment, an observer male was allowed to watch how another male from the same troop interacted with a new female before being admitted access to the pair. The striking finding was that the observer male respected the new pair bond, even if he was dominant over the rival. However, when males from different troops were used, rival males were often attacked, especially if rank differences were large. This respect of "ownership," however, was not generalizable and did not transfer to situations where two males competed over access to food.

9.3 Primate Communication in Relation to Concepts

9.3.1 *Recruiting Friends*

How do individuals benefit from their bonds, for example, when conflict suddenly arises? In many social species, victims of aggression give acoustically distinct vocalizations that are relatively context-specific (Gouzoules et al. 1984). This behavior has also been reported for chimpanzees, which give acoustically distinct screams during agonistic interactions depending on whether they are involved in a conflict as victims or aggressors (Slocombe and Zuberbühler 2005a). Victims can alter the acoustic structure of their screams depending on the severity of the experienced aggression and the composition of the nearby audience. In particular, victims of severe attacks produced screams that significantly exaggerated the true level of aggression experienced, but they did so only if there was at least one listener in the audience who matched or surpassed the aggressor in rank, demonstrating some understanding of third-party relationships (Slocombe and Zuberbühler 2007). This type of social awareness was also evident in a recent experimental study showing that chimpanzees understand who is likely to support whom during conflicts. Using playback experiments, Wittig et al. (2014) showed that, even hours after a natural aggressive interaction had occurred, subjects were still strongly affected if they heard the aggressive “waa” barks of a “friend” of the former aggressor, which was not the case if they heard the “waa” barks of other group members. Chimpanzees, in other words, are able to recruit memories of past social interactions from different sources to make inferences about current interactions. Other evidence suggests that chimpanzee victims of aggression simultaneously address two different audiences using combined sequences of screams and “waa” barks. While screams are uttered to recruit potential allies in a nearby party, “waa” barks are directly aimed at the opponent (Fedurek et al. 2015).

9.3.2 *Advertising Food*

Primates sometimes use vocalizations in ways that suggest they attempt to inform others about relevant external events. For example, when encountering food, chimpanzees and some other primates produce specific vocalizations, usually referred to as “food calls.” Acoustically, the calls resemble the species’ “greeting calls” (Laporte and Zuberbühler 2010), which indicate the psychological proximity of physically encountering a socially relevant group member and communicating to them when finding food. In chimpanzees, the acoustic structure of food calls reflects the caller’s personal assessment of the food quality as low, medium, or high. For high-preference foods, acoustic discrimination is higher, and in one captive study it was possible to ascribe different acoustic structures to specific food items (Slocombe and Zuberbühler 2006). In a follow-up study using playback experiments, chimpanzees heard recordings of other group members' high- or low-preference food calls, and this information was sufficient to elicit foraging behavior in ways that suggested the calls were meaningful to them (Slocombe and Zuberbühler 2005b).

But why would a chimpanzee be motivated to inform others about the discovery of food? The most plausible explanation is that, in chimpanzees, feeding is an inherently social activity and that individuals prefer to feed in groups. There are plausible evolutionary arguments for why this is adaptive, mainly in terms of protection against hostile neighboring groups competing for the same food tree (Wilson et al. 2014).

Another argument is that advertising food serves as an opportunity for relationship building, and there is some physiological evidence in support of this hypothesis. In chimpanzees, urinary oxytocin levels are higher after food-sharing events compared to other types of social feeding, independent of social bonds, suggesting that chimpanzees are adapted to cooperate in this context (Wittig et al. 2014). If cooperative feeding is an adaptation, then individuals should be selective in whom they inform when finding a new food source. Two lines of evidence suggest this is the case. First, chimpanzees are more likely to produce food calls if travelling in a party with friends than with other group members (Slocombe et al. 2010). Second, in playback experiments, silently feeding males, upon hearing recordings of the arrival pant hoots of a familiar group member, were more likely to reply with food calls if the arriving animal was a friend than a non-friend, especially if rank differences were large (Schel et al. 2013b).

9.3.3 Warnings About Predators

Primates have clearly definable concepts of some of their natural predators (Zuberbühler et al. 1999). The classic example comes from fieldwork on vervet monkeys that produce acoustically distinct alarm calls to different predator classes, such as snakes, terrestrial carnivores, and raptors (Struhsaker 1967). Recipients respond to these calls in ways that suggest that the calls refer to corresponding cognitive representations of the different predator classes, suggesting some form of semantic knowledge (Seyfarth et al. 1980; Chap. 8). Over the following decades, comparable findings have been reported from other primate species, including Diana monkeys (Zuberbühler et al. 1997), chacma baboons (*Papio ursinus*, Rendall 2003), ring-tailed lemurs (*Lemur catta*, Pereira and Macedonia 1991), red-fronted lemurs, white sifakas (*Eulemur fulvus rufus* and *Propithecus verreauxi,* respectively, Fichtel and Kappeler 2002), and Guereza colobus monkeys (*Colobus guereza*, Schel et al. 2010).

Male Campbell's monkey alarm calls have been particularly well studied, with a number of findings relevant to warning about predators. Three main call types were identified, with acoustically graded intermediates, that could be further modified by optional acoustic affixes (Fig. 9.1). While un-affixed calls were mainly given after

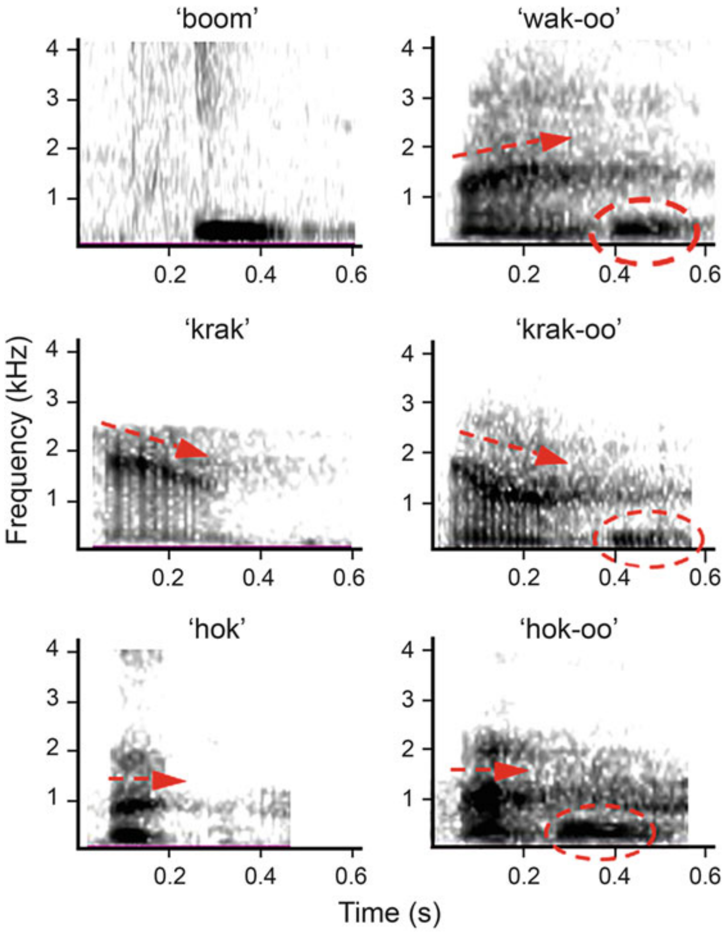


Fig. 9.1 Vocal behavior of Campbell's monkeys (*Cercopithecus campbelli*). Male Campbell's monkeys produce four basic call types to a range of external events ("boom," "krak," "hok," "wak"). Apart from booms, calls can take an optional "oo" suffix ("hok-oo," "krak-oo," "wak-oo"). Affixed calls tend to be given to unspecific disturbances, while un-affixed calls are typically predator-specific ("krak" = leopard, "hok," "wak" = eagle). Affixed calls are sometimes preceded by booms, but only in non-predatory situations (adapted from Ouattara et al. 2009a)

discovering a predator, males typically produced acoustically graded and affixed calls during non-predatory disturbances (Keenan et al. 2013). Thus, affixation appears to function to broaden the calls' meaning by transforming highly specific eagle or leopard alarms to general arboreal disturbance calls or general alert calls (Ouattara et al. 2009a).

In recent playback experiments, affixed and un-affixed calls of male Campbell's monkeys were played to groups of Diana monkeys living in the same habitat. The two species often form mixed-species groups and understand each other's vocalizations (Zuberbühler 2000b). As predicted, Diana monkeys responded significantly more strongly to un-affixed alarm calls (indicating leopard or eagle) than affixed (unspecific) alarm calls, suggesting that affixation is a communicatively important feature (Coye et al. 2015).

While female Campbell's monkeys also produce alarm calls, they are much quieter and more difficult to study under field conditions. In one study, females produced three basic alarm call types associated with specific contexts, allowing listeners to draw inferences about the type of disturbance experienced by the caller (Ouattara et al. 2009c), similar to what has been described for other primate species.

The production of acoustically distinct vocalizations is only one way by which primates communicate about external events. Alarm calls are often produced in long series, and a number of studies have focused on the information conveyed at the sequence level. A first prominent feature concerns differences in rates of call delivery. Several studies have shown that rate differences can be potentially meaningful to listeners. In Campbell's monkeys, for example, call rates of four different alarm series varied systematically with context, associated behavior, and identity of the caller. Higher call rates were generally found in predation events, especially when a predator was discovered visually, and when the caller actively attacked a predator (Lemasson et al. 2010). Call rate-based information has also been found in blue monkeys (*Cercopithecus mitis*). Here, male alarm calls consist of two types of acoustically distinct calls, and call rates were mainly affected by the distance and location of the predator (Murphy et al. 2013). A third example of call rate-based information comes from two species of colobus monkey, the King colobus (*Colobus polykomos*) and the Guereza colobus (*C. guereza*). Here, males produce sequences of loud roars to predatory eagles and leopards, but the roaring sequences differ from each other in their temporal structure (Schel et al. 2009). Interestingly, males also give roaring sequences before dawn, without any obvious external disturbance (Schel and Zuberbühler 2012a), but these call rates are slower than sequences to predators (Schel and Zuberbühler 2012b), enabling recipients to make inferences about the underlying reasons.

Another interesting feature in primate alarm calling is how acoustically different calls are assembled into sequences. In black-fronted titi monkeys (*Callicebus nigrifrons*), adults produce acoustically distinct vocalizations in response to several predator species. Call A is mainly given to raptors but also to predatory capuchin monkeys and other threats within the canopy, while call B is given to predatory or non-predatory disturbances on the ground (Casar et al. 2012a). In playback experiments, listeners responded by preferentially looking upwards when hearing raptor-related calls and towards the presumed caller when hearing terrestrial predator-related calls, with locomotor responses occurring in the expected direction (Casar et al. 2012b). Additional patterns were found at the sequence level, depending on whether the predator was a mammalian carnivore or a raptor and whether it was detected in a tree or on the ground, suggesting a capacity to encode both location and type of predatory threat within the same calling sequence (Casar et al. 2013).

In putty-nosed monkeys (*Cercopithecus nictitans martini*), males produce series of "pyows" to leopards and series of "hacks" to eagles (Eckardt and Zuberbühler

2004). In response to realistic life-sized predator models, callers produced call sequences that were highly consistent in structure, suggesting that predator class is reliably encoded at this level, irrespective of the mode of predator encounter (Arnold et al. 2008). Interestingly, some of the same call series were sometimes also given in non-predatory contexts, which was surprising because how could listeners then identify the cause of the call sequence? To address this question, females were tested with male alarm call series either alone or following additional information that simulated the occurrence of natural disturbances. Results showed that listeners were able to integrate the additional contextual information and so distinguish amongst the possible causes of the calls (Arnold and Zuberbühler 2013). Similar pragmatic abilities have also been shown in Diana monkeys responding to contextually ambiguous bird alarm calls (Zuberbühler 2000a).

One of the most surprising findings in the putty-nosed monkey alarm call system is that males sometimes produce call sequences consisting of a few “pyows” followed by a few “hacks” that are not related to any external event, but seem to function as “imperatives” to activate a group movement (Arnold and Zuberbühler 2006). In playback experiments, females discriminated between the different sequence types, mainly by staying put to “eagle” sequences, seeking additional information to “leopard” sequences, and travelling into the direction of the caller to “travel” sequences (Arnold and Zuberbühler 2008), suggesting some basic compositional abilities in these primates. The finding was also surprising because animal vocalizations are generally considered to be holistic with utterances meaning little else than the sum of their parts. In further experiments, it could be shown that monkeys perceived the “travel” sequence as a unified expression, in the sense that the component calls did not contribute individually to a combined meaning (Arnold and Zuberbühler 2012).

A third type of combinatorial system has been found in Campbell’s monkeys. As discussed earlier, adult males produce different types of loud calls (“boom,” “krak,” “hok,” “wak”; Fig. 9.1), which they combine into context-specific sequences. In an observational study, boom-introduced sequences could be associated with cohesion and travel, falling trees, neighboring groups, non-predatory animals, and unspecific predatory threats, while sequences without booms could be associated with predatory events, such as the presence of leopards and eagles (Ouattara et al. 2009b). When Diana monkeys were exposed to playbacks of boom-introduced or boom-free Campbell’s monkey alarm calls, they behaved as if the booms affected the semantic specificity of the subsequent alarm calls (Zuberbühler 2002). Specifically, playbacks of “krak” or “hok” series (male Campbell’s monkey response to leopards or eagles, respectively) triggered leopard or eagle alarm calls in male and female Diana monkeys, while playbacks of “boom”-introduced “krak” or “hok” series only triggered generalized alert responses. The interpretation was that Diana monkeys perceived the “boom” calls as a semantic modifier, which determined whether or not subsequent alarm calls referred to a specific predator type or to a general disturbance (Fig. 9.2).

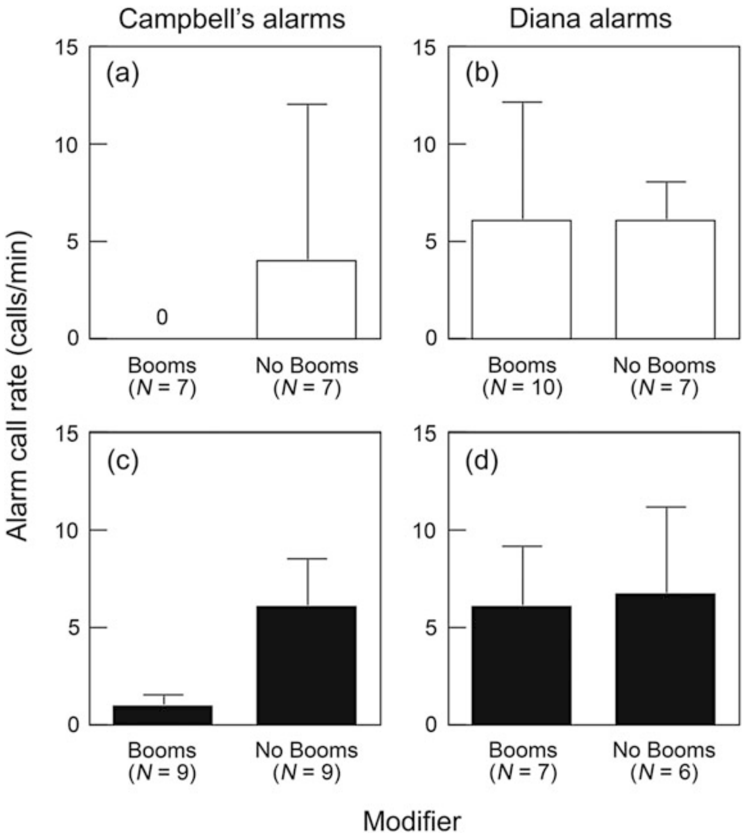


Fig. 9.2 Vocal behavior of Diana monkeys (*Cercopithecus diana*). Alarm call responses of female Diana monkeys to playbacks of male Campbell's monkey "eagle" or "leopard" alarm calls, with and without preceding "boom" calls, plus two corresponding control conditions. Playback stimuli consisted of five Campbell's monkey (**a**, **c**) or Diana monkey (**b**, **d**) alarm calls to leopards or eagles, with or without preceding booms. *Top row:* Eagle alarm calls used as stimuli; *bottom row*: leopard alarm calls used as stimuli. *Bars* and *whiskers* indicate, respectively, the median Diana monkey alarm call rates and the third quartile during the first minute after beginning of a playback. *White bars:* Eagle alarm call given in response; *black bars:* leopard alarm call given in response (adapted from Zuberbühler 2002)

9.4 Is Communicating with Regard to Concepts Learned?

The previous sections have provided an overview of how primates represent their social worlds and how these cognitive representations reveal themselves in both signalers' production of communication signals and receivers' responses. An obvious and important question then is how cognitive representations are formed, and how individuals learn to communicate them. Primates are notorious for their apparent inability to imitate sounds, very much in contrast to humans, songbirds, and other groups of mammals (Janik and Slater 1997). Various observations suggest that the limitations are due to a lack of control over the sound production mechanism that requires controlled airflow through the larynx (Lameira et al. 2014). So what are the limits of primate vocal production? A general finding here is that there is little variability within each primate species in terms of the basic types of vocalizations they can produce, but that there can be considerable acoustic variation within some of the call types. In one study, the contact calls of female Campbell's monkeys were found to vary in relation to their social bonds, independent of genetic relatedness. Social factors thus appear to influence the production of call variants in some primate calls, while the main call types remain innately determined (Lemasson et al. 2011). This line of research is still in its infancy and more progress is expected, especially in species that possess acoustically highly graded vocal systems, such as chimpanzees or bonobos.

But acquiring communication skills is also about learning to produce signals in contextually appropriate ways (see Chap. 10). Early work on vervet monkeys showed that young monkeys enter the world with predetermined notions of which classes of events warrant "eagle" alarms and which "leopard" alarms. Throughout their ontogeny, individuals then learn to apply the different call types to the biologically relevant classes of predators, probably due to socially learned predator concepts (Seyfarth and Cheney 1986). A recent study worth mentioning in this context relates to the integration of adult chimpanzees into an existing zoo colony, which has been discussed before (Schel et al. 2013a). During the integration process, the food calls of the different individuals were monitored, and the striking finding was that, after 3 years, the immigrant chimpanzees converged in their calling behavior to the patterns present in the resident group (Watson et al. 2015), further suggesting that social learning can determine how primates use their innately available vocal repertoire in response to external events (but see Fischer et al. 2015).

Ontogenetic studies with wild primates are notoriously difficult to carry out, and thus comparably little is known about the development of vocal behavior in primates (but see Chap. 10). In one study, the development of chimpanzee greeting calls was monitored. In adults, greeting is always directed at higher ranking group members, and the alpha male obtains a large share of these calls, followed by the other adult males of the group. Although infant chimpanzees produce grunts from very early on, they are not socially directed until about 2 months of age. This is followed by a period of social directedness, but calls are not given to socially relevant individuals before adolescence, showing that the acquisition of greeting behavior in chimpanzees is a long-lasting process with distinct developmental phases in which social influences by the mother and other group members are likely to play a role, most likely mediated by the formation of relevant social concepts (Laporte and Zuberbühler 2011).

Another way to study learning in primate vocal behavior is by comparing populations of the same species living in different habitats. One fruitful approach has been to compare adult vocal behavior in two different African rainforests, the Tai National Park of Cote d'Ivoire and Tiwai Island of Sierra Leone. Both habitats contain the same primate species, but Tiwai Island is leopard-free, to the effect that the local primate population grows up with a major predator missing. This ecological difference has well-documented effects on the monkeys' alarm call behavior. For example, male Diana monkeys produce the same alarm call types at both sites, but differ in how they assemble their calls into sequences. On Tiwai Island, males produced identical call sequences to leopards and general disturbances, but in Tai Forest they discriminated between the two events by producing call sequences with distinct compositions, suggesting that ontogenetic experience determined how individual calls are assembled into meaningful sequences (Stephan and Zuberbühler 2008). Similarly, female "leopard" alarm calls were extremely rare on Tiwai Island, but not in Tai Forest, whereas no differences were found in female "eagle" alarm call production (Stephan and Zuberbühler 2014).

An on-going study on Campbell's monkey vocal behavior has further revealed interesting site differences between Tiwai Island and Tai Forest populations. At both sites, males produce "krak," and "hok" calls, which can optionally be affixed with "–oo," as explained earlier (Fig. 9.1; Ouattara et al. 2009a). However, the meaning of "krak" is not identical in Tai and on Tiwai. While "krak" functions as a leopard alarm call in Tai, it is used as a general alarm call on Tiwai. One interpretation of this difference is that "krak" has the same underspecified meaning in both locations (= general alarm call), but that in Tai the meaning undergoes strengthening in that the call relates to non-aerial dangerous predators (which essentially singles out leopards), whereas on Tiwai only the un-strengthened meaning is used (Schlenker et al. 2014).

9.5 Do Primates Inform Each Other?

A final challenge in the general quest to understand primate social concepts concerns the question of whether primates perceive each other as independent mental entities with cognitive states that may differ from their own. In a series of influential experiments, Hare et al. (2000, 2001) showed that chimpanzees have some limited capacities to perceive the mental world of others, and various other studies have by and large confirmed this conclusion (see Call and Tomasello 2008 for review). Not much progress has been made, however, concerning whether primates also make active use of mental state attributions when communicating with each other. A related problem is whether primate signals go beyond mere "imperatives" (e.g., Genty and Zuberbühler 2014) and instead attempt to achieve a joint reference with the addressee ("declaratives"), a capacity based on the ability to attribute mental states to others.

For chimpanzees, several lines of research are relevant as to whether primates are capable of actively informing others based on their knowledge. First, Schel et al. (2013b) found that chimpanzee alarm calls in response to a python model were socially directed and given preferentially to the arrival of friends, not to the snake itself. More importantly, calling was associated with visual monitoring of the audience and included gaze alternations and only stopped when recipients were in safety. Second, Crockford et al. (2015) tested whether chimpanzee snake alarms direct receivers' attention to the threat by broadcasting either "alert hoos" or acoustically distinguishable “rest hoos.” Subjects showed a significant increase in search behavior after “alert hoos” compared with “rest hoos,” while monitoring the call provider and the area near the call provider, suggesting that this call type is used to actively refer recipients to relevant external events. A final study relevant in this context concerned the chimpanzees’ behavior in response to snakes in the presence of different audiences, which could be ignorant, partially informed (heard snake alarm calls), or fully informed (saw the snake). Results showed that subjects were more likely to alarm call in the presence of unaware than aware group members, suggesting that they understood something about the knowledge state of others (Crockford et al. 2012).

9.6 Summary and Future Directions

How do primates categorize their social world and how do they communicate about it? In this chapter, I have reviewed various studies, mainly from natural populations, with direct relevance to this question. Although it is not always possible to draw strong conclusions about the nature of the mental concepts available to primates, the studies reviewed here are in line with the general hypothesis that nonhuman primates categorize their social worlds in similar ways as humans. Conspecifics are not just grouped into familiar and unfamiliar individuals, but are also carriers of conceptual features, such as friends, leader, relative, or stranger. Research on monkeys has highlighted the intricacies by which conceptual information is transferred to relevant audiences, while research on great apes has highlighted the rich social worlds and high audience awareness that characterize these species. Together, these findings indicate that key psychological predispositions for the emergence of language exist in nonhuman primates.

Although considerable progress has been made by probing the minds of wild primates, this literature still pales if compared to what has been achieved in the laboratory. One of the outstanding problems in this area of research is to better understand the nature of primates’ theories of mind. Laboratory work has made great progress in pinpointing the different ways by which primates (mainly chimpanzees) show awareness of others’ mental states. For example, several experiments have shown that chimpanzees are excellent mind readers if the challenge is to understand others’ intentions (Schmelz et al. 2011) and visual perspectives (Hare et al. 2000). However, the evidence is not so good when primates need to take into account others’ knowledge or beliefs. The study by Crockford et al. (2012) provides an interesting exception by suggesting that, in situations of danger, chimpanzees take others’ knowledge or ignorance into account, but these findings need to be replicated in future work performed with tighter controls and in behavioral contexts other than danger.

Another emerging theme is that the social relations between two individuals appear to play a crucial role in how cooperative primates are. Generally, great apes are more likely to inform others about relevant events in the world if the targeted recipient is socially relevant (i.e., either high-ranking or socially close to the caller), a relevant finding for theories of the evolution of human communication. One main argument for why only humans have evolved the capacity for language is because of high motivations to collaborate and the ability to experience shared intentions (Tomasello and Carpenter 2007), a topic not well researched in wild primates. Future empirical progress can be expected by experimentally controlling the nature of social relations between individuals in order to study the effect of such manipulations on communication behavior (e.g., see Borgeaud and Bshary 2015).

In general, much of what is currently known about primate cognition comes from individuals tested in psychological laboratories (Tomasello 2014). Although the minds of these individuals are unlikely to be fundamentally different from their wild congeners, captive primates still have grown up in species-atypical environments largely structured by humans. It is therefore reasonable to expect further progress will be made in the future by studying similar research questions in natural populations. Results from future experimental fieldwork in particular are likely to provide new insights into understanding the minds of nonhuman primates. This will be accomplished by testing individuals with an ecologically relevant developmental history and experience in dealing with evolutionarily relevant situations.

Acknowledgments I am grateful to the European Research Council (PRILANG 283871) and to the Swiss National Science Foundation for their ongoing support.

References

Arnold K, Zuberbühler K (2006) Language evolution: semantic combinations in primate calls. Nature 441:303

Arnold K, Zuberbühler K (2008) Meaningful call combinations in a non-human primate. Curr Biol 18(5):R202–R203

Arnold K, Zuberbühler K (2012) Call combinations in monkeys: compositional or idiomatic expressions? Brain Lang 120(3):303–309. doi:10.1016/J.Bandl.2011.10.001

Arnold K, Zuberbühler K (2013) Female putty-nosed monkeys use experimentally altered contextual information to disambiguate the cause of male alarm calls. PLoS ONE 8(6). doi:10.1371/journal.pone.0065660

Arnold K, Pohlner Y, Zuberbühler K (2008) A forest monkey's alarm call series to predator models. Behav Ecol Sociobiol 62(4):549–559. doi:10.1007/S00265-007-0479-Y

Azevedo FAC, Carvalho LRB, Grinberg LT, Farfel JM, Ferretti REL, Leite REP, Jacob Filho W, Lent R, Herculano-Houzel S (2009) Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. J Comp Neurol 513(5):532–541. doi:10.1002/cne.21974

Barton RA (2007) Evolutionary specialization in mammalian cortical structure. J Evol Biol 20 (4):1504–1511

Barton RA, Dunbar RIM (1997) Evolution of the social brain. In: Whiten A, Byrne RW (eds) Machiavellian intelligence II. Extensions and evaluations. Cambridge University Press, Cambridge, pp. 240–263

Barton RA, Venditti C (2014) Rapid evolution of the cerebellum in humans and other great apes. Curr Biol 24(20):2440–2444. doi:10.1016/j.cub.2014.08.056

Bergman TJ, Beehner JC, Cheney DL, Seyfarth RM (2003) Hierarchical classification by rank and kinship in baboons. Science 302(5648):1234–1236

Borgeaud C, Bshary R (2015) Wild vervet monkeys trade tolerance and specific coalitionary support for grooming in experimentally induced conflicts. Curr Biol 25(22):3011–3016. doi:10.1016/j.cub.2015.10.016

Borgeaud C, van de Waal E, Bshary R (2013) Third-party ranks knowledge in wild vervet monkeys (*Chlorocebus aethiops pygerythrus*). PLoS ONE 8(3):1–4

Borgeaud C, Alvino M, van Leeuwen K, Townsend SW, Bshary R (2015) Age/sex differences in third-party rank relationship knowledge in wild vervet monkeys, *Chlorocebus aethiops pygerythrus*. Anim Behav 102:277–284. doi:10.1016/j.anbehav.2015.02.006

Call J, Tomasello M (2008) Does the chimpanzee have a theory of mind? 30 years later. Trends Cogn Sci 12(5):187–192

Candiotti A, Coye C, Ouattara K, Petit EJ, Vallet D, Zuberbühler K, Lemasson A (2015) Female bonds and kinship in forest guenons. Int J Primatol 36(2):332–352. doi:10.1007/s10764-015-9829-1

Casar C, Byrne R, Young RJ, Zuberbühler K (2012a) The alarm call system of wild black-fronted titi monkeys, *Callicebus nigrifrons*. Behav Ecol Sociobiol 66(5):653–667. doi:10.1007/S00265-011-1313-0

Casar C, Byrne RW, Hoppitt W, Young RJ, Zuberbühler K (2012b) Evidence for semantic communication in titi monkey alarm calls. Anim Behav 84(2):405–411. doi:10.1016/J.Anbehav.2012.05.010

Casar C, Zuberbühler K, Young RJ, Byrne RW (2013) Titi monkey call sequences vary with predator location and type. Biol Lett 9(5). doi:10.1098/Rsbl.2013.0535

Cheney DL, Seyfarth RM (1982) Recognition of individuals within and between groups of free ranging vervets. Am Zool 22:519–529

Coye C, Ouattara K, Zuberbühler K, Lemasson A (2015) Suffixation influences receivers' behaviour in non-human primates. Proc R Soc B Biol Sci 282(1807):50265–50265

Crockford C, Wittig RM, Mundry R, Zuberbühler K (2012) Wild chimpanzees inform ignorant group members of danger. Curr Biol 22(2):142–146. doi:10.1016/J.Cub.2011.11.053

Crockford C, Wittig RM, Langergraber K, Ziegler TE, Zuberbühler K, Deschner T (2013) Urinary oxytocin and social bonding in related and unrelated wild chimpanzees. Proc R Soc B Biol Sci 280(1755). doi:Artn 20122765, 10.1098/Rspb.2012.2765

Crockford C, Wittig RM, Zuberbühler K (2015) An intentional vocalization draws others' attention: a playback experiment with wild chimpanzees. Anim Cogn 18(3):581–591. doi:10.1007/s10071-014-0827-z

Darwin C (1871/1998) Chapter 3. Comparison of the mental powers of man and the lower animals. In: The descent of man, and selection in relation to sex. Prometheus, New York, NY, pp. 66–99

Dasser V (1988) A social concept in Java monkeys. Anim Behav 36:225–230

Dunbar RIM, Shultz S (2007) Evolution in the social brain. Science 317(5843):1344–1347. doi:10.1126/science.1145463

Eckardt W, Zuberbühler K (2004) Cooperation and competition in two forest monkeys. Behav Ecol 15(3):400–411

Emery NJ, Clayton NS (2004) The mentality of crows: convergent evolution of intelligence in corvids and apes. Science 306(5703):1903–1907

Fedurek P, Slocombe KE, Zuberbühler K (2015) Chimpanzees communicate to two different audiences during aggressive interactions. Anim Behav 110:21–28

Fichtel C, Kappeler PM (2002) Anti-predator behavior of group-living Malagasy primates: mixed evidence for a referential alarm call system. Behav Ecol Sociobiol 51(3):262–275

Fischer J, Wheeler BC, Higham JR (2015) Is there any evidence for vocal learning in chimpanzee food calls? Curr Biol 25(21):R1028–R1029. doi:10.1016/j.cub.2015.09.010

Genty E, Zuberbühler K (2014) Spatial reference in a bonobo gesture. Curr Biol 24 (14):1601–1605. doi:10.1016/j.cub.2014.05.065

Genty E, Clay Z, Hobaiter C, Zuberbühler K (2014) Multi-modal use of a socially directed call in bonobos. PLoS ONE 9(1). doi:ARTN e84738, 10.1371/journal.pone.0084738

Gouzoules S, Gouzoules H, Marler P (1984) Rhesus monkey (*Macaca mulatta*) screams: representational signalling in the recruitment of agonistic aid. Anim Behav 32(1):182–193

Grosenick L, Clement TS, Fernald RD (2007) Fish can infer social rank by observation alone. Nature 445(7126):429–432

Hare B, Call J, Agnetta B, Tomasello M (2000) Chimpanzees know what conspecifics do and do not see. Anim Behav 59(4):771–785

Hare B, Call J, Tomasello M (2001) Do chimpanzees know what conspecifics know? Anim Behav 61(1):139–151

Herbinger I, Papworth S, Boesch C, Zuberbühler K (2009) Vocal, gestural and locomotor responses of wild chimpanzees to familiar and unfamiliar intruders: a playback study. Anim Behav 78(6):1389–1396

Humphrey NK (1976) The social function of intellect. In: Bateson PPG, Hinde RA (eds) Growing points in ethology. Cambridge University Press, Cambridge, UK, pp. 303–317

Janik VM, Slater PJB (1997) Vocal learning in mammals. Adv Study Behav 26:59–99

Keenan S, Lemasson A, Zuberbühler K (2013) Graded or discrete? A quantitative analysis of Campbell's monkey alarm calls. Anim Behav 85(1):109–118. doi:10.1016/J.Anbehav.2012.10.014

Kummer H, Gotz W, Angst W (1974) Triadic differentiation: an inhibitory process protecting pair bonds in baboons. Behaviour 49(1–2):62–87

Lameira AR, Maddieson I, Zuberbühler K (2014) Primate feedstock for the evolution of consonants. Trends Cogn Sci 18(2):60–62. doi:10.1016/J.Tics.2013.10.013

Langergraber KE, Rowney C, Crockford C, Wittig R, Zuberbühler K, Vigilant L (2014) Genetic analyses suggest no immigration of adult females and their offspring into the Sonso community of chimpanzees in the Budongo Forest Reserve, Uganda. Am J Primatol 76(7):640–648. doi:10.1002/ajp.22258

Laporte MNC, Zuberbühler K (2010) Vocal greeting behaviour in wild chimpanzee females. Anim Behav 80(3):467–473

Laporte MNC, Zuberbühler K (2011) The development of a greeting signal in wild chimpanzees. Dev Sci 14(5):1220–1234. doi:10.1111/J.1467-7687.2011.01069.X

Lemasson A, Ouattara K, Bouchet H, Zuberbühler K (2010) Speed of call delivery is related to context and caller identity in Campbell's monkey males. Naturwissenschaften 97 (11):1023–1027. doi:10.1007/s00114-010-0715-6

Lemasson A, Ouattara K, Petit EJ, Zuberbühler K (2011) Social learning of vocal structure in a nonhuman primate? Bmc Evol Biol 11. doi:10.1186/1471-2148-11-362

Murphy D, Lea SEG, Zuberbühler K (2013) Male blue monkey alarm calls encode predator type and distance. Anim Behav 85(1):119–125. doi:10.1016/J.Anbehav.2012.10.015

Noe R, Laporte M (2014) Socio-spatial cognition in vervet monkeys. Anim Cogn 17(3):597–607. doi:10.1007/s10071-013-0690-3

Ouattara K, Lemasson A, Zuberbühler K (2009a) Campbell's monkeys use affixation to alter call meaning. PLoS ONE 4(11)

Ouattara K, Lemasson A, Zuberbühler K (2009b) Wild Campbell's monkeys concatenate vocalizations into context-specific call sequences. Proc Natl Acad Sci USA 106(51):22026–22031

Ouattara K, Zuberbühler K, N'goran EK, Gobert J-E, Lemasson A (2009c) The alarm call system of female Campbell's monkeys. Anim Behav 78:35–44

Palombit RA (1999) Infanticide and the evolution of pair bonds in nonhuman primates. Evol Anthrop:117–129

Pereira ME, Macedonia JM (1991) Ringtailed lemur anti-predator calls denote predator class, not response urgency. Anim Behav 41:543–544

Rendall D (2003) The affective basis of referential grunt variants in baboons. J Acoust Soc Am

Sallet J, Mars RB, Noonan MP, Andersson JL, O'Reilly JX, Jbabdi S, Croxson PL, Jenkinson M, Miller KL, Rushworth MFS (2011) Social network size affects neural circuits in macaques. Science 334(6056):697–700. doi:10.1126/science.1210027

Schel AM, Zuberbühler K (2012a) Dawn chorusing in guereza colobus monkeys. Behav Ecol Sociobiol 66(3):361–373. doi:10.1007/S00265-011-1282-3

Schel AM, Zuberbühler K (2012b) Predator and non-predator long-distance calls in Guereza colobus monkeys. Behav Process 91(1):41–49. doi:10.1016/J.Beproc.2012.05.004

Schel AM, Tranquilli S, Zuberbühler K (2009) The alarm call system of two species of black-and-white colobus monkeys (*Colobus polykomos* and *Colobus guereza*). J Comp Psychol 123 (2):136–150

Schel AM, Candiotti A, Zuberbühler K (2010) Predator-deterring alarm call sequences in Guereza colobus monkeys are meaningful to conspecifics. Anim Behav 80(5):799–808

Schel AM, Rawlings B, Claidiere N, Wilke C, Wathan J, Richardson J, Pearson S, Herrelko ES, Whiten A, Slocombe K (2013a) Network analysis of social changes in a captive chimpanzee community following the successful integration of two adult groups. Am J Primatol 75 (3):254–266. doi:10.1002/Ajp.22101

Schel AM, Townsend SW, Machanda Z, Zuberbühler K, Slocombe KE (2013b) Chimpanzee alarm call production meets key criteria for intentionality. PLoS ONE 8(10). doi:ARTN e76674, 10.1371/journal.pone.0076674

Schlenker P, Chemla E, Arnold K, Lemasson A, Ouattara K, Keenan S, Stephan C, Ryder R, Zuberbühler K (2014) Monkey semantics: two 'dialects' of Campbell's monkey alarm calls. Linguist Philos 37(6):439–501. doi:10.1007/s10988-014-9155-7

Schmelz M, Call J, Tomasello M (2011) Chimpanzees know that others make inferences. Proc Natl Acad Sci USA 108(7):3077–3079. doi:10.1073/pnas.1000469108

Seyfarth RM, Cheney DL (1982) How monkeys see the world: a review of recent research on East African vervet monkeys. In: Snowdon CT, Brown CH, Petersen MR (eds) Primate communication. Cambridge University Press, Cambridge, p. S246

Seyfarth RM, Cheney DL (1986) Vocal development in vervet monkeys. Anim Behav 34:1640–1658

Seyfarth RM, Cheney DL (2012) The evolutionary origins of friendship. Annu Rev Psychol 63:153–177. doi:10.1146/annurev-psych-120710-100337

Seyfarth RM, Cheney DL, Marler P (1980) Monkey responses to three different alarm calls: Evidence of predator classification and semantic communication. Science 210:801–803

Slocombe KE, Zuberbühler K (2005a) Agonistic screams in wild chimpanzees (*Pan troglodytes schweinfurthii*) vary as a function of social role. J Comp Psychol 119(1):67–77

Slocombe KE, Zuberbühler K (2005b) Functionally referential communication in a chimpanzee. Curr Biol 15(19):1779–1784

Slocombe KE, Zuberbühler K (2006) Food-associated calls in chimpanzees: responses to food types or food preferences? Anim Behav 72:989–999

Slocombe KE, Zuberbühler K (2007) Chimpanzees modify recruitment screams as a function of audience composition. Proc Natl Acad Sci USA 104:17228–17233

Slocombe KE, Kaller T, Turman L, Townsend S, Papworth S, Squibbs P, Zuberbühler K (2010) Production of food-associated calls in wild male chimpanzees is dependent on the composition of the audience. Behav Ecol Sociobiol:65

Stephan C, Zuberbühler K (2008) Predation increases acoustic complexity in primate alarm calls. Biol Lett 4(6):641–644. doi:10.1098/Rsbl.2008.0488

Stephan C, Zuberbühler K (2014) Predation affects alarm call usage in female Diana monkeys (*Cercopithecus diana diana*). Behav Ecol Sociobiol 68(2):321–331. doi:10.1007/S00265-013-1647-X

Struhsaker TT (1967) Auditory communication among vervet monkeys (*Cercopithecus aethiops*). In: Altmann SA (ed) Social communication among primates. University of Chicago Press, Chicago, IL, pp. 281–324

Tobias P (1971) The brain in hominid evolution. Columbia University Press, New York, NY

Tomasello M (2014) A natural history of human thinking. Harvard University Press, Cambridge, MA

Tomasello M, Carpenter M (2007) Shared intentionality. Dev Sci 10(1):121–125

Townsend SW, Slocombe KE, Thompson ME, Zuberbühler K (2007) Female-led infanticide in wild chimpanzees. Curr Biol 17(10):R355–R356

Townsend SW, Deschner T, Zuberbühler K (2008) Female chimpanzees use copulation calls flexibly to prevent social competition. Plos One 3(6):e2431

Tromp D, Meunier H, Roeder JJ (2015) Transitive inference in two lemur species (*Eulemur macaco* and *Eulemur fulvus*). Am J Primatol 77(3):338–345. doi:10.1002/ajp.22349

Watson SK, Townsend SW, Schel AM, Wilke C, Wallace EK, Cheng L, West V, Slocombe KE (2015) Vocal learning in the functionally referential food grunts of chimpanzees. Curr Biol 25 (4):495–499. doi:10.1016/j.cub.2014.12.032

Wilson ML, Hauser MD, Wrangham RW (2001) Does participation in intergroup conflict depend on numerical assessment, range location, or rank for wild chimpanzees? Anim Behav 61:1203–1216

Wilson ML, Boesch C, Fruth B, Furuichi T, Gilby IC, Hashimoto C, Hobaiter CL, Hohmann G, Itoh N, Koops K, Lloyd JN, Matsuzawa T, Mitani JC, Mjungu DC, Morgan D, Muller MN, Mundry R, Nakamura M, Pruetz J, Pusey AE, Riedel J, Sanz C, Schel AM, Simmons N, Waller M, Watts DP, White F, Wittig RM, Zuberbühler K, Wrangham RW (2014) Lethal aggression in *Pan* is better explained by adaptive strategies than human impacts. Nature 513 (7518):414–417

Wittig RM, Crockford C, Deschner T, Langergraber KE, Ziegler TE, Zuberbühler K (2014) Food sharing is linked to urinary oxytocin levels and bonding in related and unrelated wild chimpanzees. Proc R Soc B Biol Sci 281(1778). doi:Unsp 20133096, 10.1098/Rspb.2013.3096

Wolf U, Rapoport MJ, Schweizer TA (2009) Evaluating the affective component of the cerebellar cognitive affective syndrome. J Neuropsychiatry Clin Neurosci 21(3):245–253

Zuberbühler K (2000a) Causal cognition in a non-human primate: field playback experiments with Diana monkeys. Cognition 76:195–207

Zuberbühler K (2000b) Interspecific semantic communication in two forest monkeys. Proc R Soc London Ser B Biol Sci 267(1444):713–718

Zuberbühler K (2002) A syntactic rule in forest monkey communication. Anim Behav 63:293–299. doi:10.1006/Anbe.2001.1914

Zuberbühler K (2014) Experimental field studies with non-human primates. Curr Opin Neurobiol 28:150–156. doi:10.1016/j.conb.2014.07.012

Zuberbühler K, Janmaat K (2010) Foraging cognition in non-human primates. In: Platt ML, Ghazanfar AA (eds) Primate neuroethology. Oxford University Press, Oxford, pp. 64–82

Zuberbühler K, Noë R, Seyfarth RM (1997) Diana monkey long-distance calls: messages for conspecifics and predators. Anim Behav 53:589–604

Zuberbühler K, Cheney DL, Seyfarth RM (1999) Conceptual semantics in a nonhuman primate. J Comp Psychol 113(1):33–42

Chapter 10
Decisions to Communicate in Primate Ecological and Social Landscapes

Camille R. Toarmino, Vladimir Jovanovic, and Cory T. Miller

Abstract In many species, communication is anything but the effortless task it appears. While many animals communicate with apparent ease, there are serious challenges that need to be overcome for effective signal production and reception. Primates, like many species, are faced with communicative barriers in both their ecological and social landscapes. The ecological landscape is often cluttered with biotic and abiotic noise as potential sources of signal masking and degradation, while the social landscape is riddled with the nuances of the complex relationships that exist between individuals. Here, we consider how different facets of these landscapes pose challenges for communication and discuss research examining how common marmosets (*Callithrix jacchus*) overcome these obstacles. We emphasize the importance of decision-making and behavioral flexibility with regard to if and when vocalizations are produced in order to optimize effective communication. The strategies presented here are likely to exist across other primate species and reflect the broader social cognitive mechanisms within the primate Order.

C.R. Toarmino (✉)
Department of Psychology, University of California San Diego, La Jolla, CA, USA
e-mail: ctoarmin@ucsd.edu

V. Jovanovic
Neurosciences Graduate Program, University of California San Diego, La Jolla, CA, USA
e-mail: vjovanov@ucsd.edu

C.T. Miller
Department of Psychology, Neurosciences Graduate Program, University of California
San Diego, La Jolla, CA, USA
e-mail: corymiller@ucsd.edu

M.A. Bee, C.T. Miller (eds.), *Psychological Mechanisms in Animal Communication*, Animal Signals and Communication 5,
DOI 10.1007/978-3-319-48690-1_10

10.1 Introduction

Communication has traditionally been characterized as the exchange of signals between a signaler and receiver (Dawkins and Krebs 1978, 1979). A signal is produced to provide receivers with information and potentially affect subsequent behavior. An emphasis on signals has been prevalent throughout the study of animal communication systems, including in nonhuman primates. This has occurred, in part, because of the central role signals play in all facets of communication. Vocal signals convey a wealth of information about the signaler (Gerhardt 1992; Miller and Cohen 2010), such as its potential quality as a mate (Welch et al. 1998) or its individual identity (Chaps. 7 and 8). Signals can also convey specific information to signal receivers about the external world, such as the presence of food or a predator (Seyfarth et al. 1980; Evans et al. 1994; Chaps. 8 and 9). Signals have effectively coevolved with perceptual and cognitive systems in order to improve the precision with which this information can be conveyed (Miller and Bee 2012).

Vocal signals are an important mode of communication among vertebrates. However, natural environments are rarely devoid of acoustic interference (Brumm 2013; Wiley 2015). As such, the efficacy of vocal signaling is heavily affected by signalers' decisions about signal production. During communication, a signaler must decide whether to produce a signal and when this production should occur, for example, in relation to the signals of others (i.e., the call latency). Because many species produce different vocalizations based on particular contexts, signalers must also decide which of their vocal signals is most suitable for the given setting. Various ongoing social and ecological factors are likely to affect the efficacy of these signaling decisions. Therefore, the decisions made by a signaler regarding whether, when, and what to signal are crucial to ensuring that vocalizations are audible to the intended receiver(s) and that the most appropriate behavioral responses are made. Although the signal itself remains critical to communication, a signaler's decisions about whether, when, and what to signal also play important and often underappreciated roles in the inherently social behavior of communicating.

The decision-making processes integral to communication, however, are often complicated due to various obstacles that are fundamental components of the social and ecological landscapes of nonhuman primates. These landscapes thus present signalers and receivers with distinct sets of challenges for communication. In the ecological landscape, the signal itself can degrade due to acoustic interference from biotic and abiotic noise, vegetation, and weather, among other factors. Selection has acted on the acoustic structure of species' vocal signals to avoid some of the tonic sources of masking interference. For example, some vocalizations have evolved within a specific frequency range to avoid interference from calls produced by heterospecifics (Marten and Marler 1977; Marten et al. 1977; Waser and Waser 1977; Wiley and Richards 1982; Brown and Waser 1988). However, most sources of interference are not as constant and reliable. An unexpected gust of wind or the call of a conspecific produced in close proximity could significantly degrade signaling efficacy. Individual signalers, therefore, must possess a degree of behavioral flexibility that affords a capacity to decide the most suitable time to emit a vocalization. Mechanisms for deciding to produce a call or to manipulate the acoustic structure of the signal would provide a means to effectively communicate in the face of the challenges of the ecological landscape (Brumm et al. 2004; Eliades and Wang 2008; Roy et al. 2011).

The primate social landscape presents an entirely distinct set of complications for communication. In this domain, barriers to communication do not arise from degradation of the signal structure or interfering noise, but rather from how relationships between conspecific group members may affect the intended outcome of signal production. These social landscapes are filled with individuals of varying social relatedness, ranks, and allegiances (Chaps. 8 and 9). For instance, chacma baboons (*Papio ursinus*) are keenly aware of the relative rank of group members and how those relationships affect behavioral interactions (Kitchen et al. 2005). Deviating from what is considered appropriate behavior is likely to be noticed by conspecific group members (Cheney et al. 1995; Bergman et al. 2003). Ultimately, a signaler's previous interactions and experiences with the intended receivers of a signal likely influence their decision-making processes, as a particular vocalization produced in the presence of some individuals may not always yield the desired outcome. To achieve their communicative goals, primates must consider these obstacles in the social landscape when making communicative decisions.

This chapter reviews work investigating how common marmosets, a nonhuman primate that has emerged as a model organism in biomedical research (Mansfield 2003; Okano et al. 2012; Kishi et al. 2014), overcome communication problems imposed by features of their ecological and social landscapes. After a brief overview of common marmoset biology (Sect. 10.2), we separately review work on communication in the contexts of problems arising in the animal's ecological landscape (Sect. 10.3) and social landscape (Sect. 10.4). We conclude with a brief summary and a look forward to several future research directions (Sect. 10.5).

10.2 Common Marmosets

The common marmoset (*Callithrix jacchus*) is a small-bodied New World primate. Similarly to other callitrichids, marmosets live in stable social groups comprising a pair-bonded male and female, as well as their offspring (Digby and Barreto 1993). The species is entirely arboreal, inhabiting the forests of northeastern Brazil (Fig. 10.1; Hubrecht 1985; Scanlon et al. 1989; Souto et al. 2007; Rylands et al. 2009). Field studies examining common marmoset vocal communication report that this species produces a broad range of vocalizations when interacting with group members, during competitive contexts with other groups at territorial boundaries, and in response to predators (Bezerra and Souto 2008). Due to the density of vegetation in their forest habitats, visual contact between group members is often limited. Consequently, marmosets are highly voluble, engaging in frequent vocal interactions with group members. One vocal interaction that occurs often is an exchange

Fig. 10.1 A common marmoset (*Callithrix jacchus*) in the Atlantic forest at the Tapacura Ecological Field Station (Recife, Brazil)

of their species-typical contact call—the "phee" call—when individuals are visually occluded from other group members (Norcross and Newman 1993; Miller et al. 2010b). This natural vocal behavior, known as antiphonal calling, has been studied relatively extensively in laboratory conditions at both behavioral and neural levels (Norcross and Newman 1997; Miller and Wang 2006; Miller et al. 2009a, b, 2010a, b, 2015; Roy et al. 2011; Miller and Thomas 2012; Takahashi et al. 2013; Chow et al. 2015; Toarmino et al. in review). Antiphonal calling is governed by social rules, such as turn-taking (Miller and Wang 2006; Takahashi et al. 2013), in which conspecifics will alternate their phee calls to avoid interrupting each other. Differences in the temporal dynamics of these vocal exchanges are apparent across social contexts (Norcross and Newman 1997; Miller and Wang 2006) and in the presence of acoustic interference (Roy et al. 2011), suggesting that social and ecological factors affect decisions that signalers and receivers make during these interactions.

10.3 Ecological Landscape

Signalers must make key decisions about the timing and occurrence of vocal production in order to avoid the myriad sources of acoustic interference inherent to the ecological landscape. In the forests of northeastern Brazil, to which common marmosets are endemic, both biotic and abiotic noise can impact signal efficacy. Morrill et al. (2013) sought to characterize the acoustic properties of the natural forest habitat of common marmosets and the various sources of acoustic interference. Ambient recordings of the habitat acoustics in these forests revealed a significant effect of both season (wet versus dry) and time of day. A higher prevalence of acoustic interference in the frequency range of common marmoset

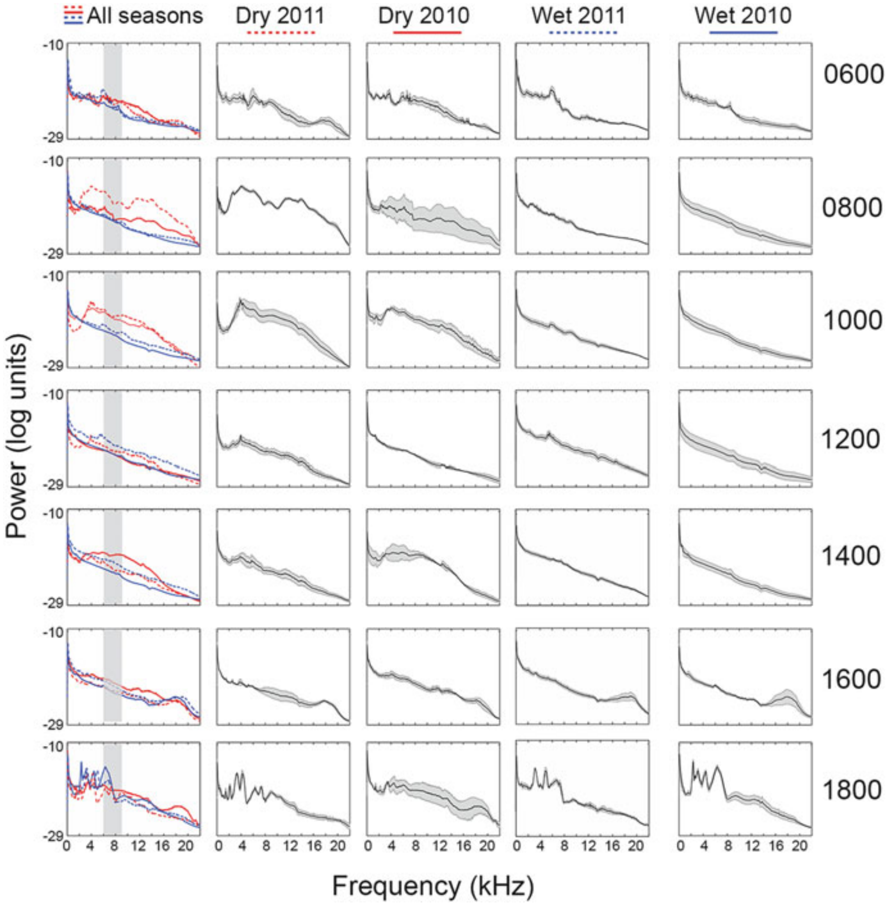


Fig. 10.2 Power spectra of ambient habitat acoustic recordings by season and time of day. *Right* margin indicates time of day. On all plots, the *x*-axis shows frequency up to 22.05 kHz and the *y*-axis shows power units, logarithmically transformed. The *leftmost* column shows the mean spectra by time of day from each season on one plot for comparison. The shaded *gray* area indicates the approximate range of the fundamental frequency of the common marmoset phee call (~6–9 kHz). The other four columns show the mean spectrum by time of day for the season indicated at the top, with each plot representing an average spectrum from 3 days of recordings. Shaded *error bars* represent ± 1 s.e.m.

calls (~6–9 kHz) was more evident during the months of the dry season compared to the wet season. Independent of the season, interference was most prevalent during the morning and late afternoon hours (Fig. 10.2). These findings suggest acoustic interference is not constant in marmoset habitats but rather changes as a function of the time of day and the season (Morrill et al. 2013). Exactly how these acoustic factors affect communication decisions in marmosets is not well known, but the variability of these ecological factors suggests that decisions of when and how to produce calls are likely critical. Notably, one behavioral study found that marmoset calling rates were fairly consistent over the course of the day (Bezerra et al. 2009), suggesting that they may be using different strategies for optimizing communication in the face of ecological challenges rather than simply avoiding particular times.

A second component of the study by Morrill et al. (2013) examined sound transmission in the same forest habitat in order to determine how this environment affected signal degradation. Broadcasts of constant and pulsed tones revealed patterns of degradation similar to those reported in previous studies (Marten and Marler 1977; Marten et al. 1977; Waser and Waser 1977; Wiley and Richards 1982; Brown and Waser 1988). Specifically, degradation occurred more rapidly at higher frequencies and when vegetation was denser during the wet season. Interestingly, similar broadcasts with phee calls revealed a different pattern. These vocalizations were more resistant to degradation, particularly during the wet season, suggesting that the acoustic properties of these calls may have been selected to improve signal transmission in this forest habitat. However, further broadcasts with phee calls in which the fundamental frequency and duration were manipulated suggested that these calls were not fully optimized for signal efficacy in these forests. Phee calls in the natural 7–8 kHz range degraded significantly more quickly than synthetically manipulated calls with a lower frequency. This suggests that while phee calls may transmit better than some sounds in this environment, they still degrade quickly over long distances in the forest. In order to maximize signal efficacy, behavioral strategies are likely needed.

Avoiding acoustic interference necessitates that signalers make quick decisions based on the dynamics of the immediate setting. While a plan to produce a particular call may precede its emission (Miller et al. 2009b), the decision of when to emit the vocalization will be affected by any co-occurring noise. A laboratory study of marmoset vocal interactions sought to investigate this process by presenting bursts of noise that varied in periodicity and predictability to pairs of visually occluded marmosets (Roy et al. 2011). In the periodic experiment, a series of alternating noise and silent periods were broadcast at 2-s, 4-s, and 8-s intervals in different conditions. Marmosets quickly adapted to these conditions and avoided initiating a call during a noise burst. Interestingly, during the 8-s condition, marmosets developed notable strategies to maintain communicative efficacy in the face of this acoustic interference. Signalers either completed an interaction within the 8-s silence period (i.e., one monkey produced a phee and the second monkey responded to that phee call), or they alternated their calls across successive 8-s silence periods (i.e., one monkey called during an 8-s silence period and the other monkey responded during the next 8-s silence period). This pattern is notable because natural antiphonal calling interactions have a latency of 3 s to 5 s between the initial call and the response call. As such, in this noisy environment, the pair of marmosets made the decision to coordinate the relative timing of their calls by either decreasing the response latency to complete the entire interaction within 8 s or by increasing the latency to alternate calls across 8-s periods of noise. Typically, phee calls that occur after a delay of about 8 s or more are not deemed antiphonal responses and are ignored (Miller et al. 2009a), but this communication rule was not strictly followed in the context of this experiment. This suggests that marmosets possess the behavioral flexibility to decide when to produce a call in order to improve signaling efficacy when faced with interference in the ecological landscape.

Communicating in a forest environment presents many challenges to signalers. Interfering noise from both biotic and abiotic sources can affect the efficacy with which vocal signals transmit throughout the forest environment. In response to more tonic sources of masking interference, selection may act on the acoustic structure of vocalizations in order to reduce degradation of the signal (Wiley and Richards 1982; Waser and Brown 1984). However, many sources of interference are periodic or less predictable. The capacity of primates to modify the acoustic structure of their calls in response to interference is limited (Miller et al. 2003; Brumm et al. 2004; Egnor et al. 2006), suggesting that the behavior of signalers may play a key role in overcoming these challenges. To combat the effects of interference, signalers must decide the most suitable time to produce a vocalization. Evidence suggests that marmosets, and other primates (Egnor et al. 2007), not only have the capacity to monitor the acoustic environment for interference, but to improve signaling efficacy by deciding to call during periods of low interference. These decisions are not only based on what is ideal for the signaler but are also made in coordination with a receiver.

10.4 Social Landscape

The complexities of social living present many challenges to communication. Signalers must make key decisions when producing a call based on the presumed effect it will have on all potential receivers, both intended and otherwise. The significance of this decision-making process is evident even in dyadic communication between two individuals. When marmosets communicate in dyads, the timing of their phee calls is important in order to avoid overlapping their calls with one another. This results in an alternation of calling, known as turn-taking (Miller and Wang 2006; Takahashi et al. 2013). This behavior suggests that marmoset vocal interactions are governed by social rules that determine when it is appropriate to produce a phee call in relation to the other individual's call. The precise timing of these exchanges is perceptually salient to marmosets. Interactive playback experiments in which the latency of a speaker to respond to the subject with a phee call was systematically manipulated showed that calls produced beyond a certain window of time (approximately 8 s) were significantly less likely to elicit a response from receivers (Miller et al. 2009a). The importance of the timing of phee calls suggests that timing likely guides communication decisions and relays important information about the signaler, perhaps regarding its willingness to communicate. Recent work showed how this decision-making process develops over ontogeny and changes as the conversations expand from dyads to networks.

A longitudinal ontogenetic study of marmoset vocal exchanges suggests that turn-taking is a learned behavior and that parents may guide its development (Chow et al. 2015). Natural vocal interactions between young marmosets, their siblings, and their parents were recorded over the first year of life. Analyses indicated that early in development, infant marmosets did not consistently follow the turn-taking pattern. Rather, infants were significantly more likely than adults to interrupt other signalers and produce inappropriate call types (i.e., non-phee calls) during these vocal exchanges. Although most signalers exhibited normal patterns of vocal behavior by 7–8 months, the developmental trajectory was context dependent. Marmosets exhibited correct turn-taking in their vocal exchanges with their mothers significantly earlier in ontogeny than with their fathers (Fig. 10.3). This suggested that young marmosets needed to learn different rhythms in their interactions with each respective parent. Learning the behavior appeared to be guided by parents, as they provided corrective feedback to their offspring. Following interruptions of their own calls, parents did not produce a response, effectively ending

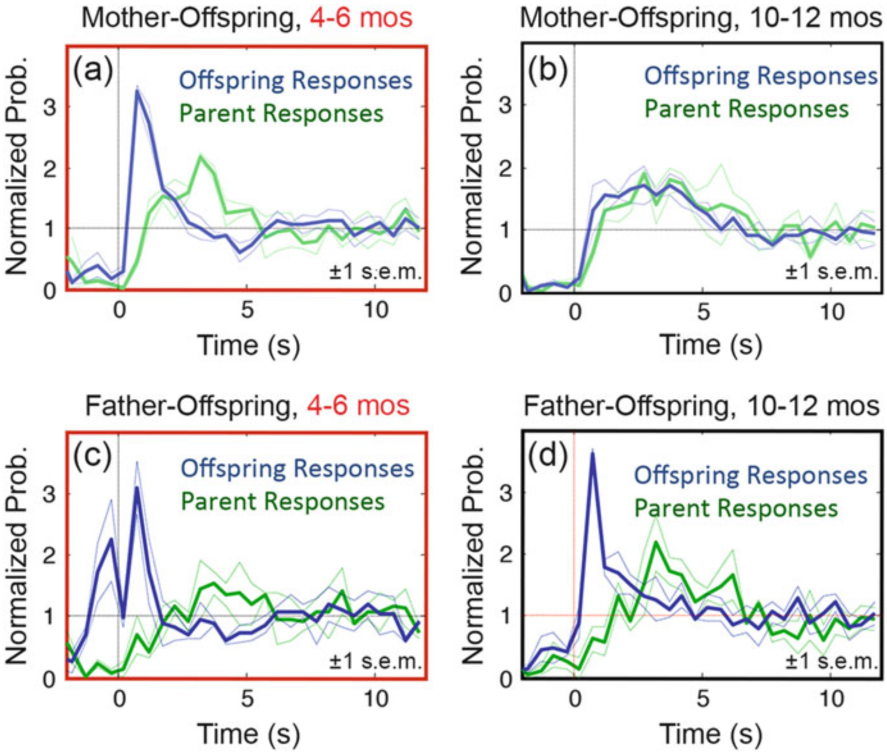


Fig. 10.3 Parent-specific patterns of conversational turn-taking. This figure plots the normalized probability (± 1 s.e.m., *lighter lines*) of the response onset for parent-offspring conversations either early (4–6 months) or late (10–12 months) in the year. (**a**) Mother-offspring conversations, early months. (**b**) Mother-offspring conversations, late months. (**c**) Father-offspring conversations, early months. (**d**) Father-offspring conversations, late months

the conversation, and were more likely to interrupt their offspring during the production of non-phee calls. Overall, these findings suggest that turn-taking is not only learned, but that the behavior's development is also driven by the signaler's understanding of the important nuances of its social landscape (i.e., who is present) and the appropriateness of its decisions about signal timing and type of signal to produce. The significance of these decisions will only increase as the landscape increases in complexity from a dyad to multiple signalers.

Due to the gregarious nature of nonhuman primate societies, communication rarely occurs in isolated dyads. Rather, communication occurs in environments with multiple individuals capable of signaling and receiving within range of each other, a hallmark of an animal communication network (McGregor 2005; Miller and Bee 2012). In these contexts, the behavior of individuals is emphasized, rather than solely the acoustics of a signal. Despite the prevalence of communication networks in primate societies, this context is not typically considered in studies of primate acoustic communication. However, a recent study has shed light on factors that influence decision-making in a marmoset communication network.

Previous research has shown that, in dyads, cues such as latency are important for engaging conspecifics in vocal exchanges (Miller et al. 2009a). But how do marmosets determine with whom to communicate when multiple individuals are available? Using a novel design, Toarmino et al. (in review) examined this question by implementing an experimentally controlled marmoset communication network consisting of "virtual monkeys" (VMs). VMs were speakers that broadcast conspecific phee calls such that each VM represented one individual conspecific. Each condition in the experiment consisted of one, two, or four VMs.

VMs varied along two categories in response to subjects' vocalizations that included the likelihood of a response (high, 80 %; low, 20 %) and the latency with which the response occurred (short, 1 s to 3 s; long, 13 s to 15 s). In this experiment, two main events could occur: subjects could either initiate a vocal exchange with a VM, or a VM could initiate a vocal exchange with the subject. When the first event occurred, that is, when the subject vocalized, one VM responded according to its likelihood and latency categories. In this way, subjects learned which VMs expressed a willingness to communicate with them (i.e., high likelihood/short latency). During the second event, one VM was selected to broadcast a spontaneous phee call during periods of silence from the subject. This was done to simulate a conspecific engaging the subject in communication. These events were important because they tested whether subjects exhibited preferences for communicating with particular VMs based on what they had learned about each VM's vocal behavior during prior events. The likelihood that subjects responded to these spontaneous calls from each VM was recorded.

This study found that, in dyads, subjects indiscriminately communicated with whatever VM was available, regardless of what kind of likelihood and latency combination it displayed (i.e., high-short, high-long, low-short, low-long). However, when there were two or four VMs with which the subject could choose to communicate, the strategy of subjects changed. Subjects exhibited a preference for VMs that were highly likely and quick to respond. Still, these cues by themselves were insufficient to elicit preferential responding from subjects. Rather, subjects' decisions to preferentially engage with a particular VM type were based on cues related to both likelihood and latency (high-short) being displayed by a VM, as well as on how other VMs in the simulated network were vocally behaving. Only when another VM was present that was unlikely and slow to respond did marmosets selectively engage with the other VM that was quick and reliable in responding. These findings suggest that the vocal behavior of multiple VMs influenced subjects' decisions to communicate. It appears that marmosets monitor their social environments and attribute certain vocal behavior to specific individuals. They selectively engage with conspecifics willing to engage with them when others who are unwilling to vocally interact are also present. This pattern of conditional preferences for engagement suggests that marmosets make judgments about the vocal behavior of conspecifics and that these judgments are relative to the immediate social context.

The types of decisions discussed here are not unique to marmosets. Other primates make similar decisions regarding whether or not to vocally interact with conspecifics. For instance, the decision to produce an alarm call or remain silent can have serious consequences for others within the group. As alarm calls function to warn conspecifics of impending danger, the decision to not produce a call in the presence of a predator can result in death. Conversely, the consequences of emitting an alarm call appear to affect the signaler and receiver in different ways, depending on the context. Tufted capuchin monkeys (*Cebus apella*), for instance, produce "hiccup" alarm calls that function to elicit an escape response from conspecifics within the vicinity of the signaler. In some cases, however, these monkeys produce alarm calls in the presence of food, rather than a predator (Wheeler 2009). Occasionally this causes nearby conspecifics to display antipredator escape reactions, leading them away from a potential food resource. In these scenarios, the decision to vocalize under false pretense benefits the signaler by increasing its chances of usurping food resources, whereas receivers miss out on food and expend energy fleeing. In some cases, wild capuchins act to inform conspecifics of a food source by producing food calls. Similar to alarm calls, they do not always use food calls appropriately. For example, wild capuchins produce food calls when other conspecifics are close by, more so than when they are alone (Gros-Louis 2004). This occurs more often when a high-ranking individual is present, as it appears to function to mitigate aggression from the individual toward the signaler. These results suggest that capuchin monkeys, like marmosets, monitor their social environment for the behavior of their conspecifics. Their decisions to communicate are flexible and can be updated to reflect the immediate social context.

10.5 Summary and Future Directions

Humans and nonhuman primates are regularly presented with the challenges of communicating in their respective ecological and social landscapes. While the ecological landscape can be cluttered with various sources of biotic and abiotic noise that degrade signal structure, appropriate behavioral decisions can serve to mitigate these effects and optimize signaling efficacy. Control over the timing of when a signal is produced seems especially relevant to solving this problem. The social landscape, on the other hand, is embedded with social nuances of the group and the relationships of conspecifics within hearing range of a signal. In this context, signalers and receivers appear to consider the relative rank of others, the relationships of those in proximity, and how previous interactions with particular individuals will affect communication. To this end, signalers must make decisions based on numerous factors in the social landscape. Their decisions to vocalize must take into account not only the effects that the signal may have on the intended receiver but also on other individuals in the vicinity.

Dealing with these environmental and social pressures has had a significant impact on the evolution of human and nonhuman primates (Seyfarth and Cheney 2014). It has been proposed that social interactions, in particular, were a driving force for the advanced cognitive abilities seen in both humans and nonhuman primates (Byrne and Whiten 1989; see also Chap. 9). Vocal behaviors are in many ways a reflection of the social behaviors that characterize primate societies. Remembering previous interactions, for instance, likely contributed to the formation of alliances that are of great importance in primate social systems (Mitani et al. 2012). The capacity of primates to overcome these challenges and find solutions to the problems of signal degradation and communicating socially likely occurred early in primate evolution and may have contributed to the evolution of the diverse range of cognitive abilities evident across the primate Order.

The use of active social signaling paradigms, such as in the interactive virtual monkey (VM) playback experiments described here, represents a potentially powerful approach for studies of primate communication and is a cornerstone for our future research (Miller et al. 2016). By implementing these experimental designs, we are able to directly engage subjects in their communicative interactions and, therefore, use these designs to examine facets of communication not possible with more traditional primate playback methods. Within these active social signaling paradigms, one can experimentally manipulate properties of the social landscape and investigate many aspects of primate communication that have so far eluded systematic experimental inquiry, such as natural social decision-making and social monitoring. Further study of these processes within the context of natural communication likely holds the key to elucidating the close relationship between primate cognition and communication and how the confluence of their underlying mechanisms provide individuals with the sophisticated social knowledge needed to navigate the intricacies of the primate social landscape. Furthermore, combining active social signaling paradigms, such as the virtual monkey experiments described here, with methods to record the activity of individual neurons in freely moving, naturally behaving monkeys (Miller et al. 2015), is another critical avenue of research needed to better understand the neural mechanisms underlying primate vocal behaviors and social interactions. The neurobiology of primate communication has been poorly studied, relative to several other taxonomic groups, but the recent development of a broad behavioral and neural toolkit provides exciting opportunities to investigate neural circuitry underlying communication in a manner not previously possible (Miller et al. 2016). The frontier of research will be in deftly combining these methods to probe the sophisticated elements of the primate social brain.

Acknowledgments This work was supported by grants from the NIH (R00 DC009007, R01 DC012087) and NSF (IDBR 1254309)

References

Bergman TJ, Beehner JC, Cheney DL, Seyfarth RM (2003) Hierarchical classification by rank and kinship in baboons. Science 302:1234–1236

Bezerra BM, Souto A (2008) Structure and usage of the vocal repertoire of *Callithrix jacchus*. Int J Primatol 29:671–701

Bezerra BM, Souto A, Borstelmann de Oliveira MA, Halsey LG (2009) Vocalizations of wild common marmosets are influenced by diurnal and ontogenetic factors. Primates 50:231–237

Brown C, Waser P (1988) Environmental influences on the structure of primate vocalizations. In: Todt D, Goedeking P, Symmes D (eds) Primate vocal communication, vol 51–68. Springer, Berlin

Brumm H (ed) (2013) Animal communication and noise. Animal signals and communication. Springer, New York

Brumm H, Voss K, Kollmer I, Todt D (2004) Acoustic communication in noise: regularion of call characteristics in a New World monkey. J Exp Biol 207:443–448

Byrne RW, Whiten A (1989) Machiavellian intelligence: social expertise and the evolution of intellect in monkeys, apes and humans. Oxford University Press, Oxford

Cheney DL, Seyfarth RM, Silk J (1995) The responses of female baboons (*Papio cynocephalus ursinus*) to anomalous social interactions: Evidence for causal reasoning? J Comp Psychol 109:134–141

Chow C, Mitchell J, Miller CT (2015) Vocal turn-taking in a nonhuman primate is learned during ontogeny. Proc R Soc B 282:210150069

Dawkins R, Krebs JR (1978) Animal signals: information or manipulation. In: Krebs JR, Davies NB (eds) Behav Ecol, vol 2. Blackwell, Oxford, pp. 282–309

Dawkins R, Krebs J (1979) Arms race between and within species. Proc R Soc B 205:489–511

Digby LJ, Barreto CE (1993) Social organization in a wild population of *Callithrix jacchus*: part 1: group composition and dynamics. Folia Primatol 61:123–134

Egnor SER, Iguina C, Hauser MD (2006) Perturbation of auditory feedback causes systematic pertubation in vocal structure in adult cotton-top tamarins. J Exp Biol 209:3652–3663

Egnor SER, Wickelgren JG, Hauser MD (2007) Tracking silence: adjusting vocal production to avoid acoustic interference. J Comp Physiol A 193:477–483

Eliades SJ, Wang X (2008) Neural substrates of vocalization feedback monitoring in primate auditory cortex. Nature 453:1102–1106

Evans CS, Evans L, Marler P (1994) On the meaning of alarm calls: functional reference in an avian vocal system. Anim Behav 45:23–38

Gerhardt HC (1992) Multiple messages in acoustic signals. Semin Neurosci 4:391–400

Gros-Louis J (2004) The function of food-associated calls in the white faced capuchin monkeys, *Cebus capucinus*, from the perspective of the signaler. Anim Behav 67:431–440

Hubrecht RC (1985) Home range size and use and territorial behavior in the common marmoset, Callithrix jacchus jacchus, at the Tapacura Field Station, Recife, Brazil. Int J Primatol 6:533–550

Kishi N, Sato K, Sasaki E, Okano H (2014) Common marmoset as a new model animal for neuroscience research and genome editing technology. Dev Growth Differ 56(1):53–62

Kitchen DM, Cheney DL, Seyfarth RM (2005) Male chacma baboons (*Papio hamadryas ursinus*) discriminate loud call contests between rivals of different relative ranks. Anim Cogn 8:1–6

Mansfield K (2003) Marmoset models commonly used in biomedical research. Comp Med 53 (4):383–392

Marten K, Marler P (1977) Sound transmission and its significance for animal vocalization. 1. Temperate habitats. Behav Ecol Sociobiol 2:271–290

Marten K, Quine DB, Marler P (1977) Sound transmission and its significance for animal vocalization. 2. Tropical habitats. Behav Ecol Sociobiol 2:291–302

McGregor PK (ed) (2005) Animal communication networks. Cambridge University Press, Cambridge

Miller CT, Bee MA (2012) Receiver psychology turns 20: is it time for a broader approach? Anim Behav 83:331–343

Miller CT, Cohen YE (2010) Vocalizations as auditory objects: behavior and neurophysiology. In: Platt M, Ghazanfar AA (eds) Primate neuroethology. Oxford University Press, New York, NY, pp. 237–255

Miller CT, Thomas AW (2012) Individual recognition during bouts of antiphonal calling in common marmosets. J Comp Physiol A 198:337–346

Miller CT, Wang X (2006) Sensory-motor interactions modulate a primate vocal behavior: antiphonal calling in common marmosets. J Comp Physiol A 192:27–38

Miller CT, Flusberg S, Hauser MD (2003) Interruptibility of cotton-top tamarin long calls: implications for vocal control. J Exp Biol 206:2629–2639

Miller CT, Beck K, Meade B, Wang X (2009a) Antiphonal call timing in marmosets is behaviorally significant: Interactive playback experiments. J Comp Physiol A 195:783–789

Miller CT, Eliades SJ, Wang X (2009b) Motor-planning for vocal production in common marmosets. Anim Behav 78:1195–1203

Miller CT, Dimauro A, Pistorio A, Hendry S, Wang X (2010a) Vocalization induced cFos expression in marmoset cortex. Front Integr Neurosci 4(128):121–115

Miller CT, Mandel K, Wang X (2010b) The communicative content of the common marmoset phee call during antiphonal calling. Am J Primatol 72:974–980

Miller CT, Thomas AW, Nummela S, de la Mothe LA (2015) Responses of primate frontal cortex neurons during natural vocal communication. J Neurophysiol 114:1158–1171

Miller CT, Freiwald W, Leopold D, Mitchell JF, Silva A, Wang X (2016) Marmosets: a neuroscientific model of human social behavior. Neuron 90:219–233

Mitani JC, Call J, Kappeler PM, Palombit RA, Silk JB (eds) (2012) The evolution of primate societies. University of Chicago, Chicago, IL

Morrill R, Thomas AW, Schiel N, Souto A, Miller CT (2013) The effect of habitat acoustics on common marmoset vocal signal transmission. Am J Primatol 75:904–916

Norcross JL, Newman JD (1993) Context and gender specific differences in the acoustic structure of common marmoset (*Callithrix jacchus*) phee calls. Am J Primatol 30:37–54

Norcross JL, Newman JD (1997) Social context affects phee call production by nonreproductive common marmosets (*Callithrix jacchus*). Am J Primatol 43:135–146

Okano H, Hikishima K, Iriki A, Sasaki E (2012) The common marmoset as a novel animal model system for biomedical and neuroscience research applications. In: Seminars in fetal and neonatal medicine, vol 6. Elsevier, pp 336–340

Roy S, Miller CT, Gottsch D, Wang X (2011) Vocal control by the common marmoset in the presence of interfering noise. J Exp Biol 214:3619–3629

Rylands AB, Coimbra-Filho AF, Mittermeier RA (2009) The systematics and distributions of the marmosets (*Callithrix*, *Callibela*, *Cebuella*, and *Mico*) and callimico (*Callimico*) (Callitrichidae, Primates). In: Ford SM, Porter LM, Davis LC (eds) The smallest anthropoids: the marmoset/callimico radiation. Springer, New York, NY, pp. 25–62

Scanlon CE, Chalmers NR, Monteiro da Cruz MAO (1989) Home range use and exploitation of gum in the marmoset *Callithrix jacchus jacchus*. Int J Primatol 10:123–136

Seyfarth RM, Cheney DL (2014) Evolution of language from social cognition. Curr Opin Neurobiol 28:5–9

Seyfarth RM, Cheney DL, Marler P (1980) Monkey responses to three different alarm calls: Evidence of predator classification and semantic communication. Science 210:801–803

Souto A, Bezerra BM, Schiel N, Huber L (2007) Saltatory search in free-living *Callithrix jacchus*: environmental and Age influences. Int J Primatol 28:881–893

Takahashi D, Narayanan D, Ghazanfar AA (2013) Coupled oscillator dynamics of vocal turn-taking in monkeys. Curr Biol 23:2162–2168

Toarmino CR, Wong L, Miller CT (in review) Social context affects decision-making in a simulated primate communication network

Waser P, Brown CH (1984) Is there a "sound window" for primate communication? Behav Ecol Sociobiol 15:73–76

Waser PM, Waser MS (1977) Experimental studies of primate vocalization: specializations for long-distance propagation. Zeitschrift fur Tierpsychologie 43:239–263

Welch AM, Semlitsch RD, Gerhardt HC (1998) Call duration as an indicator of genetic quality in male gray tree frogs. Science 280(5371):1928–1930

Wheeler BC (2009) Monkeys crying wolf? Tufted capuchin monkeys use anti-predator calls to usurp resources from conspecifics. Proc R Soc Lond Ser B Biol Sci 276:3013–3018

Wiley RH (2015) Noise matters: the evolution of communication. Harvard University Press, Cambridge, MA

Wiley RH, Richards DG (1982) Adaptations for acoustic communication in birds: sound propagation and signal detection. In: Kroodsma DE, Miller EH (eds) Acoustic communication in birds, vol 1. Academic Press, New York, NY, pp. 131–181

Chapter 11
Overcoming Sensory Uncertainty: Factors Affecting Foraging Decisions in Frog-Eating Bats

Rachel A. Page and Patricia L. Jones

Abstract Predators forage in complex environments where they must make fast, high-stakes decisions. Foraging decisions are influenced by biases in sensory perception and cognitive processing, learned and remembered information, and environmental factors such as prey availability. In this chapter, we discuss some of the factors that influence decision-making in a neotropical predatory bat species, the fringe-lipped bat, *Trachops cirrhosus*. This bat hunts frogs and insects by eavesdropping on prey-produced sounds, but its foraging decisions are also influenced by other sources of information, including echoacoustic and gustatory cues. *T. cirrhosus* quickly learns novel associations between prey cue and quality, can use social information acquired from conspecifics, and forms long-term memories of prey sounds. Research on perception and cognition in this predatory bat, all conducted with wild or wild-caught and temporarily housed individuals, has made this species one of the most well-understood, non-model systems for predator decision-making. Yet there is still much that remains unknown about how and why these predators make the foraging decisions they do.

11.1 Introduction

Optimal foraging theory predicts that individuals make foraging decisions that maximize fitness (Emlen 1966; MacArthur and Pianka 1966). But in nature, uncertainties abound, and for each foraging decision, an animal must make the best of imperfect information. Foraging animals have access to multiple sources of

---

R.A. Page (✉)
Smithsonian Tropical Research Institute, Balboa, Ancón, Republic of Panama
e-mail: pager@si.edu

P.L. Jones
Department of Ecology & Evolutionary Biology, Cornell University, Ithaca, NY, USA
e-mail: plj6@cornell.edu

M.A. Bee, C.T. Miller (eds.), *Psychological Mechanisms in Animal Communication*, Animal Signals and Communication 5,
DOI 10.1007/978-3-319-48690-1_11

information, each with costs and benefits. Prey cues can be obscured by a cacophony of background information assaulting multiple sensory modalities. Predators may also vary in their experience with different prey species and may have incomplete knowledge about which prey are palatable and which are poisonous, which are easy to capture and subdue, and which may turn the tables and attack the predator. Foraging mistakes at the least incur wasted resources of time or energy. Such mistakes can also have much greater costs, such as when errors in prey assessment result in predators mistakenly consuming toxic prey. Movement through the environment, especially the conspicuous movement involved in prey pursuit and capture, not only can give prey warning of impending attack but also can put captors in the risky position of alerting their own predators to their presence and potentially becoming prey themselves. Predator decision-making encompasses a wide array of risks, but without taking these risks, and doing so efficiently, predators lack the calories they need. The decision of when, where, and what to hunt is critical to predator survival.

The investigation of predator decision-making consists of intertwined lines of research investigating sensation, perception, and cognition. Studies of sensation and perception are concerned with identifying the sensory inputs to which a predator attends and how a predator's sensory system has evolved to increase sensitivity to certain prey cues over others. In studies of cognition, questions focus on how a predator shifts its attention between sensory cues, how it integrates and utilizes input from multiple sensory modalities, how long predators remember learned prey cues, and how these factors vary with fluctuations in social and environmental conditions. Both the sensory and cognitive components of predator decision-making impose selective consequences on predator foraging success and on prey survival.

11.1.1 Sensation, Perception, and Cognition in Predator Decision-Making

Predators use multiple sensory modalities to detect and localize prey. Sensory systems can influence foraging choices through both their separate sensitivity and tuning and through their perceptual integration. Sensory systems of animals are often restricted or tuned to be most responsive to a particular subset of the range of available stimuli (Chap. 2). A means for prey signals to be less detectable to predators is to be outside the restricted range of a predator's sensory system (Håstad et al. 2005; Stuart-Fox et al. 2008). Tuning of sensory systems therefore has important consequences for prey detection. Limitations in prey detection in one sensory system can be alleviated by the use of an additional sensory modality. The use of multiple sources of information can improve the accuracy and speed of decision-making (Rowe 1999; Roberts et al. 2007; Ward and Mehner 2010) and may be particularly beneficial in complex or uncertain environments (Rhebergen et al. 2015; Chap. 5). Multimodal stimuli can also be learned faster than stimuli in only a single sensory modality (Rowe 1999). To understand predator response to prey cues, it is, therefore, important to examine not only the response of different sensory systems to particular prey cues but additionally how these sources of sensory information are perceptually integrated.

Learning and memory also play important roles in predator decision-making. Learning abilities can be subject to natural selection due to environmental variation (Mettke-Hofmann 2014), foraging niche (Clarin et al. 2013), predation pressure (Brown and Braithwaite 2005), and social complexity (Byrne and Bates 2007). Studying the evolution of cognition in the field is a particularly challenging research area that is currently of great interest in behavioral ecology (Morand-Ferron and Quinn 2015). The majority of the research on foraging-related cognition in non-model systems has focused on birds, from tool use in New Caledonian crows (*Corvus moneduloides*, Hunt and Gray 2004) to food caching in black-capped chickadees (*Poecile atricapillus*, Pravosudov and Clayton 2002) and social learning in great tits (*Parus major*, Aplin et al. 2015). Learning enables predators to be flexible in their responses to prey cues to take advantage of temporal and geographic variation in prey availability.

11.1.2 Study System

Our system for examining the perceptual and cognitive factors influencing decision-making is the fringe-lipped bat, *Trachops cirrhosus*. *T. cirrhosus* is in the family Phyllostomidae, the leaf-nosed bats, and is found in tropical lowland forests from southern Mexico to Brazil (Cramer et al. 2001). It is a mid-sized tropical bat (~30 g, ~40-cm wingspan, Fig. 11.1) that feeds on a wide variety of prey species, including insects, frogs, lizards, and other small vertebrates (Gardner 1977; Pine and Anderson 1979; Kalko et al. 1996; Bonato and Facure 2000; Bonato et al. 2004; Rodrigues et al. 2004; Giannini and Kalko 2005). *T. cirrhosus* roosts in small mixed-sex groups in hollow trees (often cashews, *Anacardium excelsum*, Kalko et al. 1999), tunnels, caves, and culverts (Jones 1966; Handley 1976). Individuals forage in overlapping areas that average around 60 hectares (Jones et al. in review). In 1977, Merlin Tuttle was mist netting for bats at the Smithsonian Tropical Research Institute on Barro Colorado Island (BCI) in Panama and captured a *T. cirrhosus* holding a túngara frog in its mouth (Tuttle 2015). He began to investigate, comparing capture rates of *T. cirrhosus* in mist nets with and without speakers broadcasting túngara frog calls and presenting different acoustic stimuli in a flight cage. In 1980, he approached then graduate student Michael J. Ryan, who, under the mentorship of A. Stanley Rand, was studying the mating calls of túngara frogs on BCI. Together they decided to examine how *T. cirrhosus* hunts frogs. When they discovered that *T. cirrhosus* will attack speakers broadcasting túngara frog calls (Tuttle and Ryan 1981), they initiated what would become now a four-

Fig. 11.1 The fringe-lipped bat (*Trachops cirrhosus*) preys on frogs by homing in on their sexual advertisement calls. In these images, a bat approaches and consumes a calling male túngara frog (*Physalaemus pustulosus*), one of its preferred prey species. Photos courtesy Lars Hedin (*upper left*) and Alexander T. Baugh (*bottom left* and *right*)

decade-long study of sensation (Sect. 11.2), perception (Sect. 11.2), and cognition (Sect. 11.3) in this extraordinary bat species (Page et al. 2014).

11.2 Sensation and Perception

In the rainforest at night, the senses are assaulted by a myriad of diverse stimuli. Frogs and insects call to attract mates, sometimes singly, sometimes in deafening choruses. Night-blooming flowers waft intense, pungent odors. And for those who can hear ultrasound, bat and insect calls bombard the soundscape. How does a predator make sense of this cacophony and, from it, target individual prey? In a problem akin to a human attempting to attend to a single string of conversation in the boisterous confusion of a loud cocktail party (Cherry 1953; Bee and Micheyl 2008), an eavesdropping predator relies on multiple streams of sensory information to perceive and target a single prey item amidst the myriad sensory cues present in a rainforest at night.

In this section on sensation and perception, we discuss the multiple sensory inputs used by *T. cirrhosus* in their hunt for frogs. *T. cirrhosus* is now well known for locating frogs by eavesdropping on frog calls. In the bat literature, this type of prey localization is referred to as “passive listening” because it relies on prey-produced cues. Echolocation, in contrast, is often referred to as “active listening” because it is a product of bat-produced echolocation calls (Schnitzler et al. 2003). To evaluate food, many species of bats use both passive and active listening (Russo et al. 2007), as well as vision (Bell 1985), olfaction (Mikich et al. 2003), and gustation (Hristov and Conner 2005). Bats therefore have access to multiple sensory systems for perceiving potential prey. The research detailed in this section has investigated factors influencing bat responses to prey in these different sensory systems, and then how they are integrated as bats make a decision to attack.

11.2.1 Passive Listening

Passive listening is believed to be common in bats that hunt in cluttered environments (such as close to the ground or vegetation), because clutter makes locating stationary prey by echolocation nearly impossible due to the effects of backward and forward masking (Neuweiler 1989; Schnitzler and Kalko 2001; Siemers and Schnitzler 2004). Other passive listening bats attend to the broadband rustling sounds of prey moving through leaf litter (Goerlitz and Siemers 2007) or the high-frequency calls of insects such as katydids (Tuttle et al. 1985; Jones et al. 2011; Falk et al. 2015). These prey-generated cues often have high-frequency components that fall in the range of the bats' hearing, which is centered on the ultrasonic frequencies of the bats' own echolocation calls.

The demonstration by Tuttle and Ryan (1981) that *T. cirrhosus* locates frogs by their calls was such an extraordinary discovery, in part, because bats were not believed to be able to hear the low frequencies (< 5 kHz) of frog calls. It was soon after discovered that unlike most other bats, *T. cirrhosus* has a peak of auditory sensitivity below 5 kHz, the frequency range of the calls of many frog species (Fig. 11.2; Ryan et al. 1983). This additional peak in sensitivity is reflected in the neuroanatomy of this bat's ear. *T. cirrhosus* has the highest number of cochlear neurons reported for any mammal and has an additional peak of neural cochlear density not reported for any other bat species (Bruns et al. 1989). This additional peak of neural density is found in the apical portion of the cochlea (Bruns et al. 1989), the portion of the cochlea sensitive to low-frequency sounds (von Békésy 1960), suggesting auditory specialization for low-frequency sounds, such as frog calls.

Although *T. cirrhosus* eavesdrops on the calls of a number of frog (Tuttle and Ryan 1981) and katydid (Falk et al. 2015) species, the majority of work has examined the relationship between *T. cirrhosus* and the túngara frog, *Physalaemus* (=*Engystomops*) *pustulosus*. The túngara frog is a small (approximately 2 g) frog in the family Leptodactylidae that occurs throughout Middle America (Fig. 11.1). Male túngara frogs can produce two types of call: simple and complex (Fig. 11.3; Chap. 4). Both simple and complex calls contain a frequency modulated sweep or "whine," consisting of a fundamental frequency that sweeps from about 0.9 kHz to 0.4 kHz and is about 300 ms in duration. The whine has several harmonics, with an average dominant frequency of about 700 Hz. Simple túngara frog calls consist of a

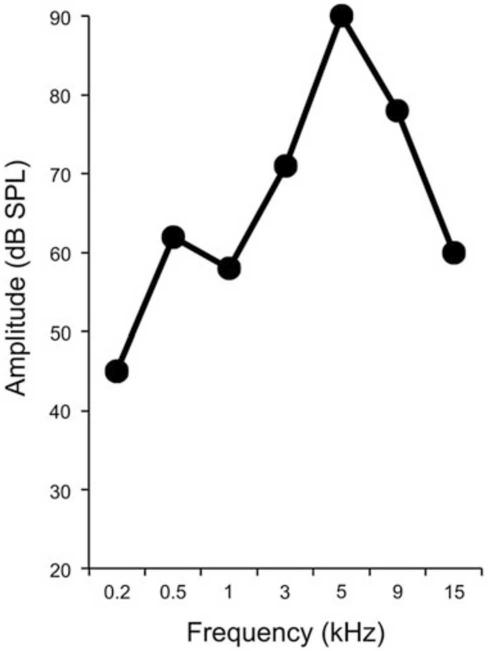


Fig. 11.2 Behavioral audiogram of a fringe-lipped bat (*Trachops cirrhosus*). Points depict the threshold amplitudes required to elicit an ear-twitching response from a perched bat in response to pure tones. Relatively lower threshold amplitudes represent greater auditory sensitivity. Note the increase in auditory sensitivity as frequencies drop below 5 kHz. Modified from Ryan et al. (1983) and used with permission

whine alone. Complex calls consist of the whine plus one to seven broadband suffixes termed “chucks.” Chucks have a dominant frequency of about 2500 Hz and a duration of 35 ms. Except in very rare cases (Ryan et al. 2015), chucks are never produced alone; they are always produced right after a whine. Both female túngara frogs (Rand and Ryan 1981; Gridi-Papp et al. 2006) and frog-eating bats (Ryan et al. 1982; Fugère et al. 2015; Akre et al. 2011) strongly prefer complex calls to simple ones, thus exerting conflicting selective pressures on the calling male.

When a group of túngara frogs are all calling together, the cacophony is such that to a human listener, it is difficult to distinguish which males are making simple calls and which are making complex calls. Locating an individual male in a loud chorus is a perceptual and cognitive challenge for female frogs (Bee and Micheyl 2008; Bee 2015; Chap. 4) and eavesdropping predators alike. A study by Jones et al. (2013a) investigated how bats respond to the different components of the túngara frog call. In particular, it examined whether bats orient toward one particular component of the call and whether they attend to the order in which the two syllables occur (in nature, chucks always follow whines). This study built off of similar research questions conducted with female túngara frogs (Farris et al. 2002, 2005; Chap. 4). Female túngara frogs strongly prefer complex calls (Gridi-Papp et al. 2006) and will not approach chucks played alone without a whine. However, if a whine is broadcast in the vicinity of a speaker playing a chuck, the females will orient toward the chuck, even with quite large spatial separations between the two call components (up to 135°, Farris et al. 2002). Female túngara frogs therefore appear to require the whine component of the call to initiate phonotaxis, indicating

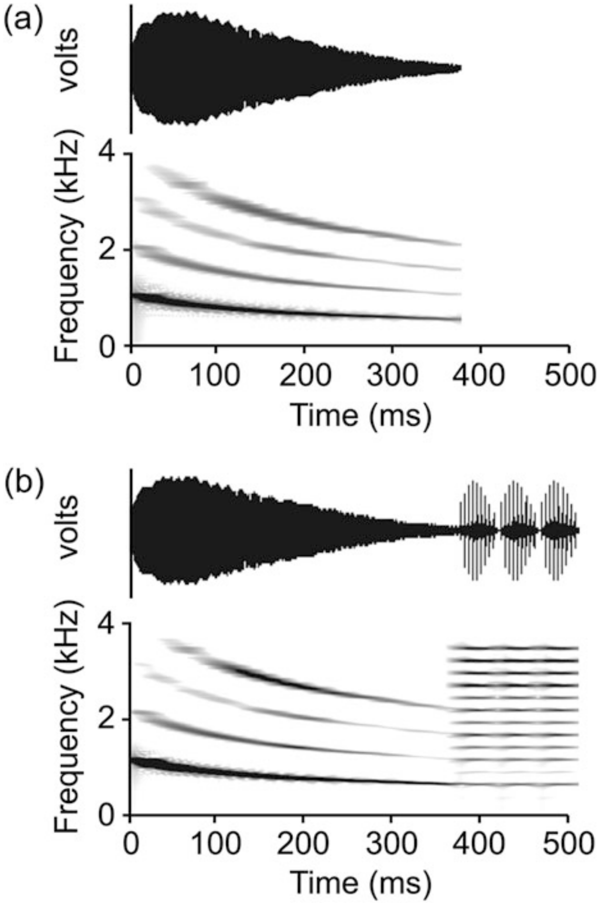


Fig. 11.3 Advertisement call of the túngara frog (*Physalaemus pustulosus*). Waveforms (*top*) and spectrograms (*bottom*) of (**a**) a simple call with no chucks and (**b**) a complex call with three chucks. Modified from Fugère et al. (2015) and used with permission

auditory grouping of these two call components, but then they preferentially approach the chuck even when it is spatially separated from the whine or broadcast in reversed temporal order such that it proceeds the whine (Farris et al. 2005).

T. cirrhosus also exhibits phonotaxis to male túngara frog calls, but bats conduct this behavior under different selective pressures and from a very divergent evolutionary starting point than female túngara frogs. Bats respond dramatically differently to isolated call components. Unlike túngara frogs, bats respond to the chuck component of the call when it is played alone, and they preferentially approach the whine over the chuck when the two components are spatially separated. If the chuck is played before the whine (the reverse of the natural order), bat approaches to the chuck increase, indicating an effect of temporal sequence (Jones et al. 2013a). These differences in responses between female túngara frogs and fringe-lipped bats highlight the different factors that weigh into decisions by these different receivers. For example, bats may be under strong selection to respond to túngara frog calls as quickly as possible so they catch their prey before it stops calling or escapes, as túngara frogs often do when they detect an approaching bat (Tuttle et al. 1982). A general preference for the first component of the call might increase the speed at which bats can make a decision to attack. The duration of a signal increases its detectability (Campbell 1963), which may explain why bats preferentially approach the whine over the chuck alone. It is important to note, however, that the whine is only preferred over the chuck *alone*. Complex calls (whines plus chucks) are still strongly preferred over simple calls (whines alone; Fugère et al. 2015). Jones et al. (2013a) showed that bats are particularly responsive to the first part of the call (in nature, this is the whine for both simple and complex calls), and bats clearly prefer the whine alone to the chuck alone, perhaps because of its duration or because the whine is a highly recognizable component of the túngara frog call. A given male facultatively makes either the whine alone or whines plus chucks. So why do bats prefer complex calls to simple ones?

Four non-exclusive hypotheses have been investigated for the preference for complex calls in *T. cirrhosus*. The first two hypotheses are based on the idea that complex calls indicate something about the prey. Bats could prefer complex calls because they signify larger males with better body condition, that is, more substantial meals. Field recordings of calling túngara frogs, however, reveal that there is no correlation between túngara frog call complexity and body length, mass, or condition (Bernal et al. 2007). A second possibility is that call complexity is an indicator of male density: complex calls indicate high-density patches of prey. This is indeed the case. The number of male túngara frogs calling within 1 m of a particular frog is correlated both with the proportion of complex calls that a male produces and with the average number of chucks he makes per call (Bernal et al. 2007). Bats may, therefore, preferentially approach complex calls because they are indicators of high prey density.

The next two hypotheses are centered around preference for call complexity as a product of bat sensory and perceptual processing. It has long been hypothesized that males that produce complex calls are easier for bats to localize because the calls are longer in duration (Campbell 1963) or because of the acoustic properties of the chuck. The chuck is short (approximately 35 ms) and has a broadband structure with a fast onset and offset, acoustic properties that should make it easier to localize (Marler 1955) than the whine. Phonotaxis experiments in a flight cage broadcasted túngara frog calls from speakers underneath screens covered in leaf litter (Page and Ryan 2008). When the bats landed on the screen, distance from the landing place to the speaker was compared for simple versus complex calls. Under most conditions, bats showed no difference in localization ability when approaching simple versus complex calls. Bats did localize complex calls better under three conditions: when (1) there was background noise present, (2) the calls were only broadcast before the bats began to approach and then shut off during approach, and (3) when the calls were only broadcast before approach and there were obstacles (hanging wooden dowels) between the perched bat and the speakers. If the calls were broadcast continuously with or without obstacles present, however, there was no improved localization (Page and Ryan 2008). The acoustic properties of the complex call that improve localization may, therefore, be only a partial explanation for bats' preference for the complex call.

The fourth hypothesis for bat preference for complex calls is that the auditory reception and processing of *T. cirrhosus* may be particularly stimulated by properties of the complex call, producing a perceptual bias for call complexity. A perceptual bias occurs when sensory and cognitive systems are biased (e.g., due to neural or chemical pathways) to be more sensitive to particular stimuli (Endler and Basolo 1998; Frame and Servedio 2012; Ryan and Cummings 2013). An example of perceptual bias is the responsiveness of mammalian auditory and cognitive systems to nonlinear sounds such as screams (Blumstein and Récapet 2009). Three different experiments have examined whether there may be a perceptual bias for complex calls in fringe-lipped bats. First, flight cage experiments presented bats with simple calls modified to contain the acoustic properties of a complex call (longer duration, greater energy, increased amplitude modulation; Fugère et al. 2015). If perceptual bias for acoustic parameters explained the preference for complex calls, then these modified calls should be equally preferred to complex calls and preferred over unmodified whines. The study also included a set of stimuli in which the acoustic parameters were exaggerated beyond those of the complex call, with the hypothesis that these stimuli should be preferred over the complex call. No modified whine was found to be more attractive than the unmodified whine, with the exception of one of the exaggerated calls modified to be as long as a complex call with six chucks (a call that is very uncommon in nature: Bernal et al. 2007; Fugère et al. 2015). This study indicates that it is possible that preference for complex calls is due to a perceptual bias for longer stimuli, but this is in need of further support.

An additional result from Fugère et al. (2015) was that bats strongly preferred the unmodified whine when paired with a modified whine that had most of its call energy in higher frequencies (in the third harmonic: over 2 kHz). Unmodified whines have most of their energy in the fundamental frequency (below 1 kHz). The bats' strong preference for unmodified whines when the alternative consisted of whines with their low frequencies removed makes sense in light of the behavioral audiogram of *T. cirrhosus*. Less sound energy is necessary to evoke a behavioral response in these bats at frequencies below 1 kHz than at frequencies above 2 kHz (Fig. 11.2; Ryan et al. 1983). The modified whines with more energy in these higher frequencies than in the lower frequencies probably sounded quieter to the bats than the unmodified whines did. As frog-eating bats prefer louder signals to quieter ones (Tuttle and Ryan 1981), their preference for unmodified whines over high-frequency modified whines is intuitive to us, as these modified calls, though matched in amplitude, likely sounded fainter to them than the unmodified calls. It is interesting to note that the dominant frequency of the túngara frog's chuck (approximately 2.5 kHz) is higher than the sonic frequencies to which *T. cirrhosus* is most sensitive. Did túngara frogs evolve away from lower-frequency chucks to reduce predation pressure from bats? An interesting avenue of research would be to compare chuck frequencies in populations of túngara frogs that vary in their degree of bat predation, as has been done in other systems, for example, in Trinidadian guppies that show increased intensity of nuptial coloration in populations under lower predation pressure (Endler 1986). Additionally, it would be interesting to compare the auditory sensitivity of bats across populations that vary in available prey. While this population comparison has not been conducted with auditory sensitivity, the two experiments discussed next did compared behavioral responses to prey calls across populations with different availability of complex calls.

Túngara frogs are not present in Amazonian Ecuador, but their sister species, Peter's dwarf frog, *Physalaemus petersi*, occurs there. As in túngara frogs, these frogs have a two-part call that consists of a frequency-modulated whine (the simple call) that can be facultatively followed by secondary component called a "squawk" (producing a complex call). Curiously, in some populations of *P. petersi*, males produce complex calls, while in other nearby populations, they do not (Boul and Ryan 2004). Trillo et al. (2013) broadcast simple and complex *P. petersi* calls from speakers in the forest and monitored bat approaches with infrared video in two locations in Amazonian Ecuador, one with complex calling *P. petersi* and the other with only simple calling *P. petersi*. Playback experiments demonstrated that bats prefer complex calls in both locations: in the area in which the frogs produce complex calls and in the area in which they produce only simple calls. It is unclear why bats prefer the calls of frogs making squawks in a population where frogs only make simple calls. It is possible that this is evidence for a perceptual bias for call complexity. It could also be that this population of frogs used to make complex calls in the past and the bat preference for complex calls is genetically controlled and is maintained. Studies from Panama show small home ranges for *T. cirrhosus* (Kalko et al. 1999), but no such tracking data is available for the Ecuadorian *T. cirrhosus*. Another possibility is that the bats fly between these frog populations and thus experience both call types. Further study is necessary to resolve these intriguing results.

A second study adds yet another layer to the story. Jones et al. (2014) investigated a population of *T. cirrhosus* at the La Selva Biological Station in Costa Rica, where *Physalaemus* is completely absent. *T. cirrhosus* in this population did not show phonotaxis toward either simple or complex túngara frog calls. The discrepancy between Trillo et al. (2013) and Jones et al. (2014) could have a number of explanations. It is possible that familiarity with simple calls predisposes bats to be more responsive to complex calls, but it is also possible that response to prey calls is genetically determined and diverges between populations that differ in prey availability.

The preference for complex calls in *T. cirrhosus*, therefore, remains a subject of debate. Future studies could fruitfully examine the responses of completely naïve, lab-reared juveniles to simple and complex túngara frog calls to test for the role of learning. Neurophysiology studies to examine how these different signals stimulate auditory neurons and parts of the brain would also make an important contribution. While passive listening for frog calls presents a number of interesting questions on sensory and perceptual processing, it is just one of the sensory systems used by bats to locate prey. The use of other sensory modalities in foraging decisions is discussed next.

11.2.2 Active Listening

The dominant sensory modality of *T. cirrhosus* when hunting frogs clearly seems to be passive listening. *T. cirrhosus* can successfully capture prey with access only to prey-produced sounds such as frog calls. However, these bats echolocate throughout the hunting approach, and studies now show that active listening using echolocation facilitates both prey localization and prey discrimination (Page et al. 2012; Halfwerk et al. 2014a). *T. cirrhosus* echolocates with short (less than 1 ms), multiharmonic, downward frequency sweeps, which range from 100 kHz to 50 kHz, with most of the call energy at 75 kHz (Barclay et al. 1981; Surlykke et al. 2013). These echolocation calls are very similar to those of other gleaning bat species in the neotropical family of leaf-nosed bats, Phyllostomidae (Falk et al. 2015). *T. cirrhosus* produces echolocation calls for orientation in space and while approaching prey (Barclay et al. 1981; Surlykke et al. 2013). Recent research has also emphasized their importance for determining prey size and for locating prey (Page et al. 2012; Halfwerk et al. 2014a). Experimental evidence even suggests *T. cirrhosus* can use echolocation cues alone to find prey in simple, uncluttered environments (Page et al. in preparation).

A túngara frog call has several by-products stemming from the production of the acoustic signal that offer additional sensory information not only to conspecifics but also to predatory bats. Like many other frog species, the túngara frog has a conspicuous vocal sac that allows it to recycle air, shuttling it back and forth between the vocal sac and the lungs. This dynamically inflating and deflating sac results in a multimodal display by the male frog (Taylor et al. 2008; Taylor and Ryan 2013; Chap. 4). Controlled experiments with a robotic frog (Fig. 11.4) have revealed that both female frogs and bats attend to the dynamically inflating vocal

Fig. 11.4 Real and robotic versions of the túngara frog (*Physalaemus pustulosus*). Photographs in the *upper row* are of an actual calling male shown in different views and states of calling. Photographs in the *bottom row* are of a robotic túngara frog. The robotic frog has a dynamically inflating vocal sac that can be synchronized with its call (see Chap. 4). Modified from Taylor et al. (2008) and used with permission

sac. Coupling motion from this sac with the acoustic signal (the frog call) increases the attractiveness of the call to both receivers, but via different sensory modalities. Female túngara frogs perceive the vocal sac using vision (Taylor et al. 2008). Bat responses to the acoustic signal, however, do not change when offered visual cues, but increase when echolocation is available (Halfwerk et al. 2014a), indicating that bats' perception of the vocal sac is with echolocation, not vision. This is particularly interesting because selective pressures from predators and mates on the same trait are thus mediated by different sensory modalities. Changes to the visual environment will affect sensory access to the multimodal display for females but not for bats, while changes to the echoacoustic environment will affect sensory access for bats but not for female frogs. Thus, environmental fluctuations could create different selective pressures on the mating signal given the differences in sensory access to the signal by mates and predators.

One of the factors that change the environment for a foraging bat is not only the density of frog choruses but also the diversity of calling frogs, as many tropical choruses contain multiple species calling at once. This complexity of sound could make the use of sensory modalities other than passive listening greatly advantageous. The use of multimodal cues is often assumed to improve signal detection and localization amidst background noise. For example, there are a number of frog species in which males call from fast-flowing streams and have evolved additional visual displays such as foot flagging ("semaphoring"), potentially because it makes their signals more salient in a noisy acoustic environment (Hödl and Amézquita 2001). If a multimodal cue makes a signal more salient for a female receiver, it may also make the signal more detectable to an eavesdropping predator. This has been demonstrated for *T. cirrhosus*: bats show more accurate angles of attack when a multimodal cue (robotic frog with inflating vocal sac) is available (Rhebergen et al. 2015). When faced with locating prey amidst heterospecific chorus noise, bats show more directional attacks on the inflating vocal sac model when there are increasing numbers of speakers broadcasting the calls of the heterospecific hourglass treefrog, *Dendropsophus ebraccatus*, which is frequently found calling in mixed species choruses with túngara frogs (Rhebergen et al. 2015). In preference tests, *T. cirrhosus* strongly prefers multimodal prey cues to unimodal ones in background noise, and both reduces the latencies of its attacks and increases its echolocation activity in response to multimodal cues in the presence of background noise (Gomes et al. 2016). The multimodality of the túngara frog call, therefore, does appear to improve localization, directionality, and attack latency for an eavesdropping predator when confronted with increased acoustic background complexity.

In addition to the dynamically inflating vocal sac, there is another signal by-product that males cannot avoid when producing their acoustic signal. Male túngara frogs call while floating on the surface of small pools of water. As they call the movement of their bodies generates water ripples that propagate through the pool (Fig. 11.5a). Frog-eating bats are more likely to attack model frogs with water ripples emanating from their calling location than model frogs with no associated ripples (Fig. 11.5b–d). This is only the case, however, when the pool is clear of leaf

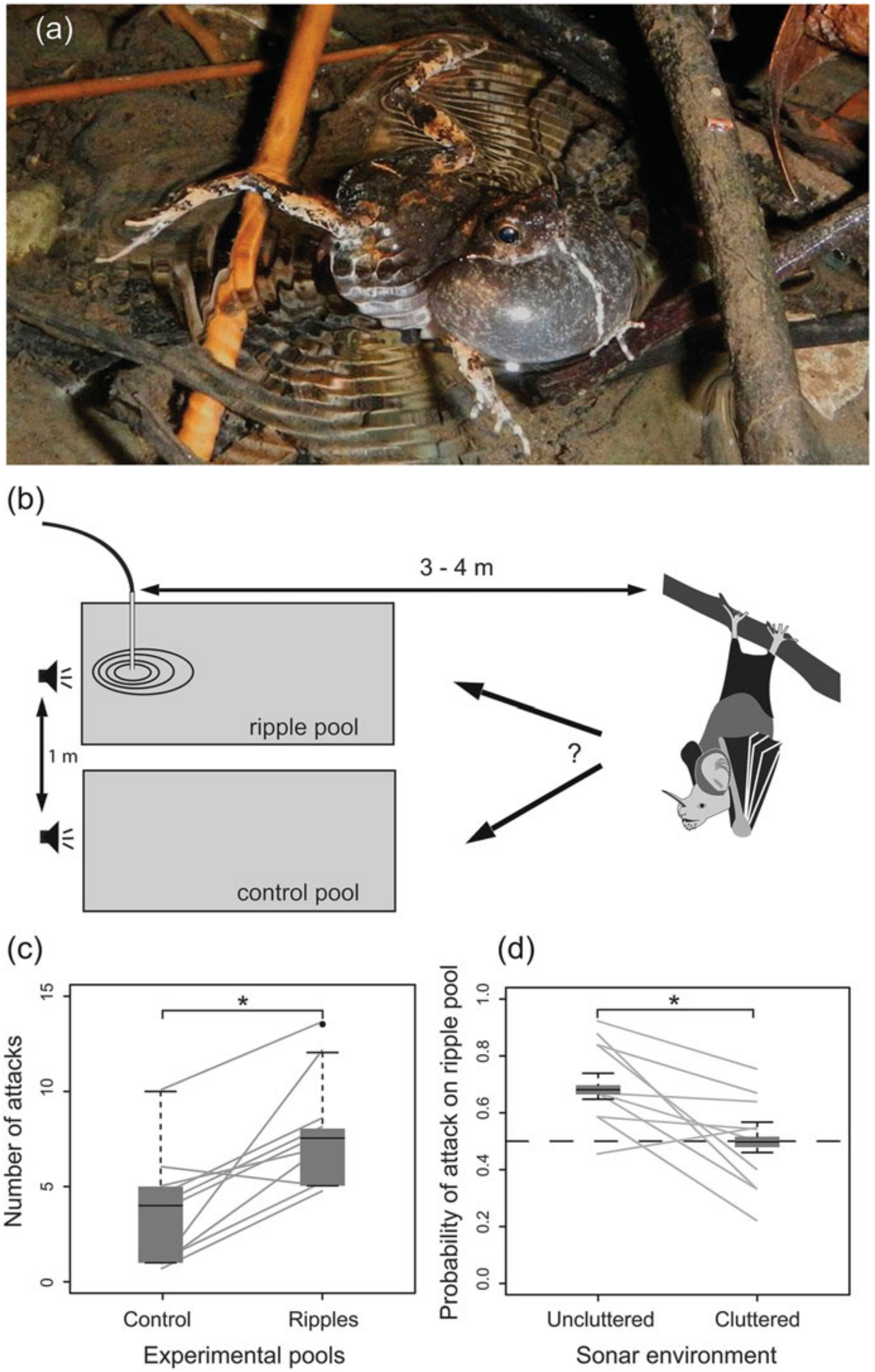


Fig. 11.5 Fringe-lipped bat (*Trachops cirrhosus*) uses echolocation to detect an unavoidable by-product of signaling by male túngara frogs (*Physalaemus pustulosus*). (**a**) A calling male

litter and therefore bats have echoacoustic access to the water ripples (Halfwerk et al. 2014b). The detectability of ripples may be an explanation for why túngara frogs are so often found calling from hidden locations underneath leaves and branches. What is so intriguing about the use of water ripples to locate calling frogs is that the ripples propagate into the environment, leaving a trace of where the frog was even if it has stopped calling. Male frogs use the water ripples in order to judge the distance that they are from another calling male in the same pond. Males generally increase their call rate when they are exposed to the call of another male paired with water ripples (Halfwerk et al. 2014b). This unintended by-product of a multimodal signal, therefore, has costs and benefits for the signaling frog and produces another cue that bats can use when locating a target to attack.

11.2.3 Chemoreception

We have discussed the importance of passive listening for prey-generated cues and echolocation to bat foraging decisions. Research has demonstrated that vision does not appear to be an important component of these foraging decisions, at least in approaches to túngara frogs (Halfwerk et al. 2014a). No research to our knowledge has investigated the role of olfaction in prey detection or discrimination with *T. cirrhosus*, although it is known to be very important for other phyllostomid bats, particularly fruit-eating species (Korine and Kalko 2005). Gustatory cues, in contrast, have been shown to be important to prey assessment in *T. cirrhosus* (Page et al. 2012). The role of taste in *T. cirrhosus* is perhaps not surprising given that many anuran species are poisonous. The cane toad, *Rhinella marina*, for example, has toxic parotoid secretions and, if consumed, is lethal to animals much larger than *T. cirrhosus* (Chen and Kovarikova 1967; Bagrov et al. 1993). Possibly to cope with anuran toxins, *T. cirrhosus* appears to have very unusual salivary glands. A study of the submandibular salivary glands of 38 genera of bats revealed that only three bat species, *T. cirrhosus*, *Megaderma lyra* (the greater false vampire bat), and *Megaderma spasma* (the lesser false vampire bat), have submandibular salivary glands containing large, follicle-like structures (Phillips et al. 1987). Later, a fourth species, *Cardioderma cor* (the heart-nosed bat), was found to possess these unusual salivary glands as well (Tandler et al. 1996). All four of these species (*T. cirrhosus*,

---

Fig. 11.5 (continued) túngara frog generates prominent ripples on the water surface while calling. (**b–d**) Bats preferentially approach the call of a túngara frog broadcast near a pool with ripples over an identical call broadcast near a pool of still water. (**b**) Schematic diagram of the experimental setup. (**c**) Results from a two-alternative choice test showing the number of attacks directed toward the ripple pool over the control pool. (**d**) Probability of attack on the ripple pool depends on environmental conditions. When both pools were covered with a layer of leaves (cluttered environment), the bats' preference for ripples disappeared. Graphs in **c** and **d** depict box plots of model estimates and individual lines. Photo in **a** courtesy of Adam Dunn; figures in **b-d** modified from Halfwerk et al. (2014b) and used with permission

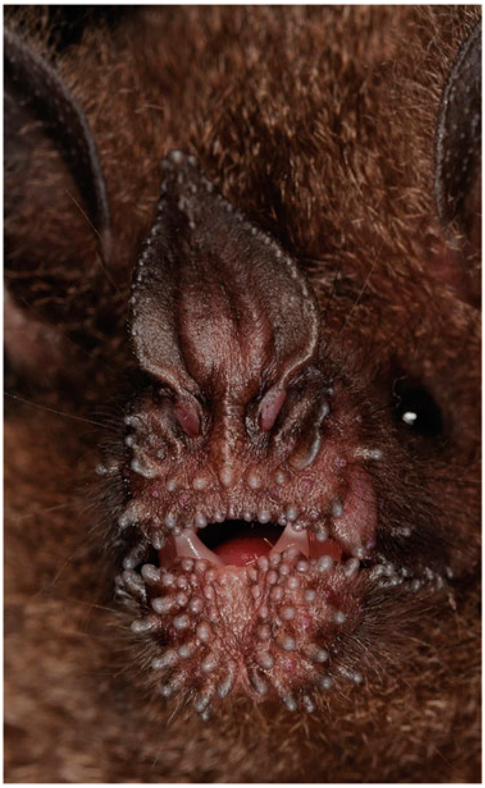


Fig. 11.6 The chin and lips of fringe-lipped bats (*Trachops cirrhosus*) are covered in distinctive tubercles, giving the species its common name. Experiments demonstrate that *T. cirrhosus* can use chemical cues to assess the palatability of its prey (Page et al. 2012). While the tubercles are hypothesized to play a role in this chemo-assessment, enabling the bat to rapidly determine the toxicity of a frog or toad just by brushing its skin, there is no evidence to support this hypothesis to date. The role of these tubercles remains a mystery and is the subject of ongoing research. Photo courtesy Marco Tschapka

C. cor, *M. lyra*, and *M. spasma*, respectively, from Latin America, East Africa, and two from Southeast Asia) feed on frogs. It has been hypothesized that the independent evolution of these unique salivary glands could be an adaptation for frog consumption. The saliva may neutralize toxins in the skin of frogs and toads, perhaps allowing bats to prey on less palatable anuran species (Phillips et al. 1987; Tandler et al. 1996).

One of the mysteries about *T. cirrhosus*, likely also related to chemoreception, is the distinctive tubercles on this bat's chin and lips (Fig. 11.6). These tubercles give this bat its common name, the fringe-lipped bat, and set it apart from other species. But to date, the function of these tubercles remains unknown. It has been hypothesized that the tubercles allow bats to rapidly assess the palatability of a frog or toad, just by brushing its tubercles to the skin of the prey prior to capture (Miller 1907). However, detailed observations with high-speed video of interactions with palatable and poisonous prey so far show no evidence of the hypothesized brushing behavior (Page et al. 2012). The role of these tubercles is the subject of ongoing investigation. Whether the mechanism is through the tubercles on the chin of *T. cirrhosus* or (more likely) through taste receptors in the mouth, bats use chemoreception in prey assessment (Page et al. 2012). When edible túngara frogs are coated with secretions from the parotoid glands of toads, bats will attack the frogs but reject them after contact (Page et al. 2012). This study highlights not only the use of chemoreception by *T. cirrhosus* but also the way that multimodal perception of prey cues can be integrated sequentially to influence foraging decisions.

11.2.4 Sequential Assessment of Prey Cues

As we have discussed in this section on perception, while passive listening is likely the dominant method used for prey detection by *T. cirrhosus*, foraging decisions are based on multiple cues that enable bats to correct potential errors and alter hunting decisions during, and even after, attack. Cues from different sensory modalities travel at different speeds and across different distances. For bats hunting frogs, the cue that travels the farthest is the male frog's advertisement call. This call likely serves as an acoustic beacon to the eavesdropping bat; it captures the bat's attention and is used by the bat to home in on its prey. As the bat approaches its target, it can recruit additional senses for prey assessment. Echolocation, for example, may be used to detect size, vocal sac movement, or environmental perturbations. When the bat approaches closer yet, olfactory cues may become available, and when it establishes physical contact with the prey, it has the possibility to use tactile and gustatory cues for the final assessment. Thus, different sensory cues are associated with the different stages of the hunting approach, and each offers an opportunity to correct mistakes in prey assessment made using another sensory modality (Fig. 11.7). It is possible that this sequential assessment of prey using multiple sensory cues is what allows frog-eating bats to be as flexible as they are in their foraging decisions, even when foraging amidst highly toxic and size-inappropriate prey.

Experimental evidence supports the idea that bats indeed use multiple sensory cues to sequentially refine their prey decisions. When offered conflicting sensory cues—a speaker broadcasting the calls of a palatable túngara frog, topped with a poisonous toad roughly the same size as a túngara frog (the leaf litter toad, *Rhinella alata*) or one much larger (the cane toad, *R. marina*)—bats will approach but veer away from the large cane toad, rejecting it before contact, but will attack the small leaf litter toad, rejecting it only after contact (Fig. 11.8a; Page et al. 2012), suggesting size-based discrimination from a distance, likely by echolocation (Halfwerk et al. 2014a). Likewise, the chemical compounds in cane toad parotoid secretions are not volatile and are secreted only upon contact (Toledo and Jared 1995), so it is unlikely that olfactory cues from these toxins alerted the bats to this toad's distastefulness before contact. It is possible that the bats smell other compounds on the frogs upon close approach. The most likely explanation for the pre-capture rejection of the large cane toad, however, is the bat's use of echolocation as it approaches the prey target. *T. cirrhosus* emits echolocation calls throughout the hunting approach (Barclay et al. 1981; Surlykke et al. 2013). It is very likely that bats approaching a prey item that is far too large for it to handle (and one that may pose a threat and could even consume the bat if offered the opportunity) use

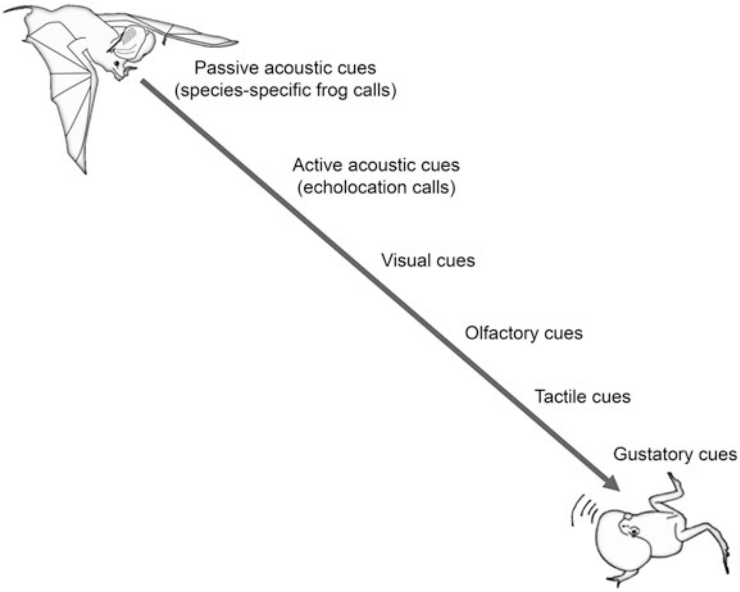


Fig. 11.7 Hypothesized sequential sensory cue use by the fringe-lipped bat (*Trachops cirrhosus*) during a hunting approach. The frog advertisement call serves as an acoustic beacon to the bat, traveling further than the other sensory cues. As the bat approaches, it can recruit other sensory cues to reassess its decision to attack the prey. Drawings courtesy Kristina Schlegel

information from their sonar signals to reject prey prior to contact (Fig. 11.8a). When bats were offered prey that differed only in chemical cues but not in size (túngara frogs rubbed either with parotoid secretions from the two toad species or with a túngara frog as a control), no frogs were rejected prior to contact. All frogs were captured, but those coated in toad secretions were rejected after the bat had come into physical contact with the prey (Fig. 11.8b). Together, these results suggest a strong reliance on sequential assessment of different prey cues. If assessment mistakes are made at one sensory level, there is the possibility for correction at another, which is perhaps key in enabling this bat to respond to prey cues as flexibly as it does.

Research with humans has shown how information from different sensory systems is weighted through experiments that present subjects with conflicting sensory information (Ernst and Bülthoff 2004). Further research on how different sensory components are weighted in foraging decisions using behavioral experiments will be important for understanding multimodal decision-making. Future research on *T. cirrhosus* will also hopefully investigate the neurobiology of sensory integration. This complexity in perception of prey cues has made research on *T. cirrhosus* groundbreaking in its insights into decision-making in a non-model system. At the same time, the extraordinary, and charismatic, learning abilities of this predator have captured the attention of cognitive ecologists. In the next section,

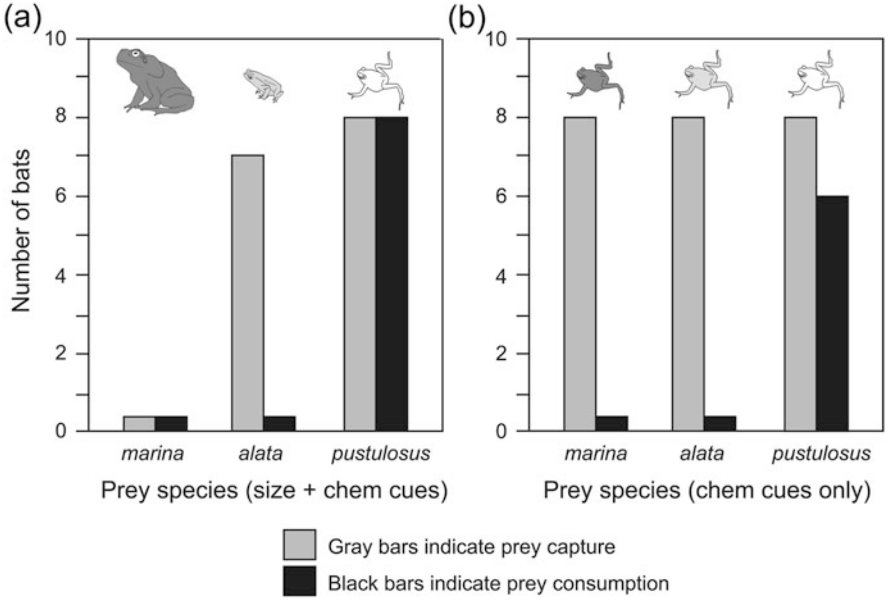


Fig. 11.8 Multimodal assessment of prey by the fringe-lipped bat (*Trachops cirrhosus*). (**a**) When both size and chemical cues are available, bats reject large poisonous toads (*Rhinella marina*) before capture, but capture and then reject small poisonous toads (*R. alata*). Palatable túngara frogs (*Physalaemus pustulosus*) are consistently captured and consumed. (**b**) When prey varies in toxicity but not in size, bats make all prey-rejection decisions post-capture. Túngara frogs rubbed in *R. marina* and *R. alata* toxins are captured and released; túngara frogs rubbed with other túngara frogs are captured and consumed ($N = 8$). Modified from Page et al. (2012) and used with permission

we discuss how learning and memory also influence foraging choices, likely overlaid on the complex integration of sensory perception.

11.3 Cognition

While research on the role of learning in foraging decisions has historically been conducted with birds, there is a developing literature on the role of learning and memory in bat foraging and social behavior. Variation in learning abilities has been demonstrated across different bat species (Clarin et al. 2013), and there is evidence for learning of group-distinctive vocalizations (Boughman 1997) and social learning of food scents (Ratcliffe and ter Hofstede 2005; O'Mara et al. 2014; Ramakers et al. 2016). *T. cirrhosus*, however, in its learning of species-specific prey cues such as frog calls (Page and Ryan 2005), remains a rare example of the study of cognition in wild-caught bats.

11.3.1 Individual Learning

Given the perils associated with prey toxicity, one might predict *T. cirrhosus* to exhibit extreme caution when making foraging decisions, with little exploration of novel prey and limited flexibility with known prey. However, there is strong evidence that even though *T. cirrhosus* forages among prey that are both highly poisonous and large enough to consume them (Ibáñez et al. 1999), these bats are extremely flexible in their hunting behavior (Page and Ryan 2005). They exhibit exploratory behavior in response to new stimuli, will generalize their response from the calls of known species to the calls of species they have never heard before (Ryan and Tuttle 1983), and can rapidly learn novel prey cues by observing conspecifics (Page and Ryan 2006; Jones et al. 2013b). They can also use individual experience to reverse established foraging patterns given new information on prey quality (Page and Ryan 2005).

There is substantial evidence that *T. cirrhosus* acquires an individual repertoire of signals they recognize as palatable prey over the course of their lifetime. When wild adult bats are brought into captivity and played frog calls from a speaker, they will readily attack the calls of palatable túngara frogs but will ignore poisonous toad calls (Tuttle and Ryan 1981). However, if a túngara frog call is gradually faded into a toad call using sound editing software, such that each sequential rewarded stimulus sounds less like a frog and more like a toad, within a night, bats will attack speakers playing poisonous toad calls to get food rewards (Page and Ryan 2005). This remarkable flexibility works the other way as well; if túngara frog calls are played repeatedly with no food rewards on the speaker, bats will eventually stop responding to túngara frog calls, although extinguishing the response to a previously positively rewarded cue takes longer than learning a new association (Page and Ryan 2005). The bats are so flexible, in fact, that their ability to lean acoustic stimuli is not limited to frog calls. They can also be quickly trained using stimuli as diverse as Bob Marley songs (R. A. Page unpublished data) and cell phone ringtones (Jones et al. 2013b). This ability to learn new associations very quickly may allow *T. cirrhosus* to develop repertoires of prey signals to which they are responsive and to shift those repertoires with changing environmental conditions such as seasonal changes in the prey species available (Jones et al. 2014).

11.3.2 Social Learning

Foraging decisions can be influenced not only by individual learning but also by acquiring information from conspecifics or social learning. Social learning is widespread in animals, potentially because it allows individuals to acquire information about the environment without incurring the potential risks of trial-and-error learning (Galef and Giraldeau 2001). Social learning of foraging information has been demonstrated for a number of bat species. Among tropical frugivorous bats, short-tailed fruit bats (*Carollia perspicillata*) learn novel food scents from demonstrators in the roost (Ratcliffe and ter Hofstede 2005), and tent-making bats (*Uroderma bilobatum*) not only learn food scents from conspecifics at the roost but specifically learn them from breath and not from odors on fur (O'Mara et al. 2014). European greater mouse-eared bats (*Myotis myotis*) can learn to associate LED lights with the presence of food and learn this task faster when they can interact with a knowledgeable demonstrator bat (Clarin et al. 2014). Similarly, North American big brown bats (*Eptesicus fuscus*) learn to capture tethered mealworms when allowed to forage alongside demonstrators, with observers and demonstrators flying close together as demonstrators made feeding buzzes in the final stages of attack (Wright et al. 2011). Given the communal roosting and foraging behavior of many bat species, social information may be easily accessible, and for bats that exploit temporary resources, a crucial source of information (Cvikel et al. 2015).

Frog-eating bats roost in small groups in culverts (Handley 1976), caves (Jones 1966), and tree cavities (Kalko et al. 1999; Jones et al. in review). Not only do they roost together, studies mounting proximity sensors to these bats have shown that they also associate at foraging sites, such as small ponds (Ripperger et al. 2016). Given the extraordinary diversity of potential prey and their seasonal variation in prey calling behavior, social learning from roost mates could be an advantageous way for frog-eating bats to acquire information about the environment. Social learning of associations between novel prey cues and food quality has been demonstrated in *T. cirrhosus*. If a naïve wild-caught bat is put in a flight cage with a bat that was previously trained to approach toad calls, the naïve bat will approach the toad call in an average of five playback trials (Page and Ryan 2006). If, in contrast, a bat is alone in the flight cage with toad calls playing and food rewards on the speaker, the naïve bats predominantly do not approach the toad calls in 100 trials. Similarly, if two naïve bats are in a flight cage together with these same acoustic stimuli and food rewards, they do not learn the toad calls any faster than a naïve bat alone, eliminating social facilitation as a potential explanation (Page and Ryan 2006). This study shows that *T. cirrhosus* is a fast social learner, providing evidence for another potential mechanism for reducing the costs of approaching novel prey.

Social learning, however, can have drawbacks in terms of misinformation, incomplete information, outdated information, or other costs associated with interacting with conspecifics, such as increased competition (Giraldeau et al. 2002; Laland 2004). The costs and benefits of social information have led researchers to predict that animals should use social information selectively, favoring high-quality (but expensive) individually acquired information under many circumstances (Laland 2004). One of the predicted social learning strategies is that animals should "copy when dissatisfied" (Laland 2004). This has been demonstrated in Norway rats (*Rattus norvegicus*) in which individuals that are fed low-quality diets are more likely to use social information to learn novel food cues than individuals fed high-quality diets (Galef et al. 2008). Similarly, bumblebees (*Bombus terrestris*) foraging on flowers with low sucrose concentration

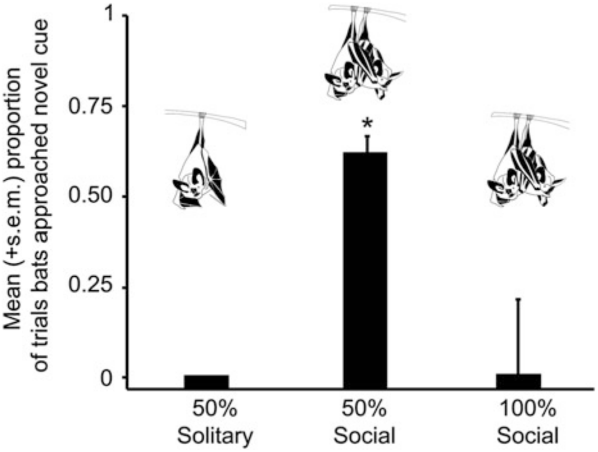


Fig. 11.9 The fringe-lipped bat (*Trachops cirrhosus*) uses social information to learn about signals produced by potential food items when their own information is unreliable. Approaches bats made to a novel ringtone depended on the reward schedule of the ringtone to which they had been trained (50 % or 100 %) and on whether or not they had social information about the novel ringtone (Social or Solitary). Bats whose own ringtone was rewarded 50 % of the time and that had social information about the novel ringtone made more approaches than bats that did not have access to social information or bats that had reliable individual information. Modified from Jones et al. (2013b) and used with permission; bat drawings courtesy Damond Kyllo

rewards are more likely to use social information to learn alternative flower colors than bees foraging on colors associated with high-concentration sucrose rewards (Jones et al. 2015). This fairly simple social learning strategy is likely widespread in animals and could be a tool many animals use when deciding whether to use social information.

To test whether frog-eating bats are using the social learning strategy of “copy when dissatisfied,” individual bats were trained to respond to one of two cell phone ringtones by fading a túngara frog call into the ringtone as with the toad experiments (Page and Ryan 2006). Bats were then divided into three treatments that differed both in the reward schedule of the ringtone to which the bat was individually trained and in the presence of a tutor bat approaching the alternative ringtone (Fig. 11.9). Bats that only received food rewards on 50 % of the presentations of their trained ringtone were more likely to approach the alternative ringtone, but only when there was a tutor bat demonstrating it (50 % Social). When there was no tutor, bats did not learn the alternative ringtone on their own, but rather continued to forage on their 50 % rewarded ringtone (50 % Solitary). When bats received food rewards for approaching the stimulus they had been trained to 100 % of the time, however, they ignored the social information about the alternative resource (100 % Social, Fig. 11.9, Jones et al. 2013b). Frog-eating bats therefore use social information to learn novel prey stimuli, but do so only under certain circumstances, weighing the costs and benefits of the sources of information in their foraging decisions.

It is unknown how important social learning may be in the wild. *T. cirrhosus* roosts together, and it is not uncommon to catch two or more adult *T. cirrhosus* in the same net at the same time when using a túngara frog chorus playback as bait (P. L. Jones and R. A. Page unpublished data). This would suggest that multiple adult *T. cirrhosus* could be foraging together and therefore have exposure to social information. Proximity sensors mounted on free-flying *T. cirrhosus* show that roost mates, and in particular what seem to be mothers and their pups, associate at foraging sites in the wild (Ripperger et al. 2016). It is possible that the close association during lactation, potentially extending several months post-weaning, may be a critical time period in which there is extensive social learning of prey cues between mothers and their young.

11.3.3 Memory

Not only do frog-eating bats quickly learn novel acoustic stimuli, but they also remember these stimuli for a long time. This may aid bats in hunting ephemeral resources that are only available for parts of the year (Jones et al. 2014). Long-term memory for learned associations between artificial cues and food has been demonstrated for greater mouse-eared bats, in which one individual after a year in the wild remembered the association between an LED light and available food (Clarin et al. 2014). Individuals of *T. cirrhosus* that were trained to fly to cell phone ringtones in the social learning experiment described above were then released back into the wild. Bats captured up to 4 years later still attacked speakers broadcasting cell phone ringtones when brought back into captivity, whereas bats that had not previously been trained to ringtones did not (Dixon et al. in preparation). This experiment demonstrates the extremely long retention of associations between prey and prey-related cues in the wild, which may be necessary for bats that forage relying on ephemeral species-specific prey signals.

11.4 Summary and Future Directions

Given its histological and neuroanatomical adaptations, including unique salivary glands and unusually low-frequency hearing suggesting specialization on frogs as prey, it is perhaps surprising to find that *T. cirrhosus* has a wide diet breadth, hunting many taxa in addition to frogs (Bonato et al. 2004), and is not stereotyped, but rather highly flexible, in its foraging behavior (Page and Ryan 2005). The studies reviewed in this chapter examined in detail the interaction between this predator's cognitive flexibility and its sensory and perceptual abilities and how they shape the foraging decisions it makes. Examination of the preference of

T. cirrhosus for complex túngara frog calls has led to discoveries about the evolution of sexual advertisement signals from the eavesdropper perspective, in terms of predator localization accuracy, population variation in predator response, auditory processing of different call components, and the role of learning in the preference for signal variants. The túngara frog is only one of a myriad of potential prey species in the Neotropics. The studies described here detail how this eavesdropper is able to be very flexible in its responses to prey calls by updating acoustic information with echoacoustic and gustatory cues as it approaches potential prey, enabling bats to avoid potentially lethal mistakes. Locating a single prey item amidst the cacophony of a frog chorus remains a challenge, however. The use of multimodal components of the frog call, particularly the inflation and deflation of the frog vocal sac and the water ripples generated during calling, enables bats to make more accurate attacks. Finally, the extraordinary flexibility, social learning, and long-term memory abilities of this bat species make it an exciting system in which to study the benefits of cognition to foraging performance.

We see two areas in need of future research. First there is a great need for understanding how animals integrate information from different sensory systems at the behavioral and neural levels. This should be addressed with behavioral experiments to disentangle when bats rely more or less heavily on different sensory systems. These experiments should especially look for cases in which the behavioral responses seem maladaptive, as these can be particularly informative about biases in sensory, perceptual, and cognitive processes. In combination, there is a great need for the integration of neurophysiology with behavioral experiments in order to understand how patterns of neural firing produce the observed behaviors.

The second area in need of research is understanding what the bats are doing in the wild. Studying the behavior of flying animals in dark, dense rainforest is impossible without the use of technological tools. Recent developments in GPS tracking, remote sensing, automated PIT tag readers, and thermal and infrared video are creating new opportunities for studying the behavior of bats in the wild. Knowledge of decision-making in this bat species would greatly benefit from more detailed data on habitat use, roosting behavior, social interactions, and long-term monitoring of individuals. Examining constraints on sensation, perception, and cognition in the wild is crucial to understanding how bats make foraging decisions in the structurally and acoustically complex environment of the biodiverse neotropical rainforest.

Acknowledgments We thank Mike Ryan, Merlin Tuttle, Ryan Taylor, Laurel Symes, Cory Miller, and Mark Bee for their helpful feedback on this chapter. We are also deeply indebted to the Smithsonian Tropical Research Institute (STRI). The body of work that has emerged over the past four decades on the sensory and cognitive ecology of the frog-eating bat would not have been possible without STRI's ongoing support.

References

Akre KL, Farris HE, Lea AM, Page RA, Ryan MJ (2011) Signal perception in frogs and bats and the evolution of mating signals. Science 333:751–752

Aplin LM, Farine DR, Morand-Ferron J, Cockburn A, Thornton A, Sheldon BC (2015) Experimentally induced innovations lead to persistent culture via conformity in wild birds. Nature 518(7540):538–541

Bagrov AY, Roukoyatkina NI, Fedorova OV, Pinaev AG, Ukhanova MV (1993) Digitalis-like and vasoconstrictor effects of endogenous digoxin-like factor(s) from the venom of *Bufo marinus* toad. Eur J Pharmacol 234(2–3):165–172

Barclay RMR, Fenton B, Tuttle MD, Ryan MJ (1981) Echolocation calls produced by *Trachops cirrhosus* (Chiroptera: Phyllostomatidae) while hunting for frogs. Can J Zool 59:750–753

Bee MA (2015) Treefrogs as animal models for research on auditory scene analysis and the cocktail party problem. Int J Psychophysiol 95(2):216–237

Bee MA, Micheyl C (2008) The cocktail party problem: what is it? How can it be solved? And why should animal behaviorists study it? J Comp Psychol 122:235–251

Bell GP (1985) The sensory basis of prey location by the California leaf-nosed bat *Macrotus californicus* (Chiroptera: Phyllostomidae). Behav Ecol Sociobiol 16:343–347

Bernal XE, Page RA, Rand AS, Ryan MJ (2007) Cues for eavesdroppers: do frog calls indicate prey density and quality? Am Nat 169(3):409–415

Blumstein DT, Récapet C (2009) The sound of arousal: the addition of novel non-linearities increases responsiveness in marmot alarm calls. Ethology 115(11):1074–1081. doi:10.1111/j.1439-0310.2009.01691.x

Bonato V, Facure KG (2000) Bat predation by the fringe-lipped bat, *Trachops cirrhosus* (Phyllostomidae, Chiroptera). Mammalia 64(2):241–243

Bonato V, Faure KG, Uieda W (2004) Food habits of bats of subfamily Vampyrinae in Brazil. J Mammal 85(4):708–713

Boughman JW (1997) Greater spear-nosed bats give group-distinctive calls. Behav Ecol Sociobiol 40(1):61–70. doi:10.1007/s002650050316

Boul KE, Ryan MJ (2004) Population variation of complex advertisement calls in *Physalaemus petersi* and comparative laryngeal morphology. Copeia 3:624–631

Brown C, Braithwaite VA (2005) Effects of predation pressure on the cognitive ability of the poeciliid *Brachyraphis episcopi*. Behav Ecol 16:482–487

Bruns V, Burda H, Ryan MJ (1989) Ear morphology of the frog-eating bat (*Trachops cirrhosus*, Family: Phyllostomidae): apparent specializations for low-frequency hearing. J Morphol 199 (1):103–118

Byrne RW, Bates LA (2007) Sociality, evolution and cognition. Curr Biol 17(16):R714–R723

Campbell RA (1963) Detection of a noise signal of varying duration. J Acoust Soc Am 35:1732–1737

Chen KK, Kovarikova A (1967) Pharmacology and toxicology of toad venom. J Pharm Sci 56 (12):1535–1541

Cherry E (1953) Some experiments on the recognition of speech, with one and with two ears. J Acoust Soc Am 25:975–979

Clarin T, Ruczynski I, Page RA, Siemers BM (2013) Foraging ecology predicts learning performance in insectivorous bats. PloS ONE 8:e64823

Clarin TMA, Borissov I, Page RA, Ratcliffe JM, Siemers BM (2014) Social learning within and across species: information transfer in mouse-eared bats. Canad J Zool 92:129–139. doi:10.1139/cjz-2013-0211

Cramer MJ, Willig MR, Jones C (2001) *Trachops cirrhosus*. Mamm Species Am Soc Mammal 656:1–6

Cvikel N, Egert Berg K, Levin E, Hurme E, Borissov I, Boonman A, Amichai E, Yovel Y (2015) Bats aggregate to improve prey search but might be impaired when their density becomes too high. Curr Biol 25(2):206–211. doi:10.1016/j.cub.2014.11.010

Dixon MM, Jones PL, Meneses S, Page RA (in preparation) Long-term memory for a novel task in wild frog-eating bats

Emlen JM (1966) The role of time and energy in food preference. Am Nat 100(916):611–617. doi:10.2307/2459299

Endler JA (1986) Natural selection in the wild. Princeton University Press, Princeton, NJ

Endler JA, Basolo AL (1998) Sensory ecology, receiver biases and sexual selection. Trends Ecol Evol 13:415–420

Ernst MO, Bülthoff HH (2004) Merging the senses into a robust percept. Trends Cogn Sci 8 (4):162–169

Falk JJ, Hofstede HM, Jones PL, Dixon MM, Faure PA, Kalko EKV, Page RA (2015) Sensory-based niche partitioning in a multiple predator–multiple prey community. Proc R Soc B Biol Sci 282:20150520

Farris HF, Rand AS, Ryan MJ (2002) The effects of spatially separated call components on phonotaxis in túngara frogs: evidence for auditory grouping. Brain Behav Evol 60:181–188

Farris HE, Rand AS, Ryan MJ (2005) The effects of time, space and spectrum on auditory grouping in túngara frogs. J Comp Physiol A Neuroethol Sens Neural Behav Physiol 191 (12):1173–1183

Frame AM, Servedio MR (2012) The evolution of preference strength under sensory bias: a role for indirect selection? Ecol Evol 2(7):1572–1583

Fugère V, O'Mara MT, Page RA (2015) Perceptual bias does not explain preference for prey call adornment in the frog-eating bat. Behav Ecol Sociobiol:1–12. doi:10.1007/s00265-015-1949-2

Galef BG Jr, Giraldeau L-A (2001) Social influences on foraging in vertebrates: causal mechanisms and adaptive functions. Anim Behav 61(1):3–15

Galef BG, Dudley KE, Whisken EE (2008) Social learning of food preferences in 'dissatisfied' and 'uncertain' Norway rats. Anim Behav 75:631e637

Gardner AL (1977) Feeding habits. In: Baker RJ, Jones JK, Carter DC (eds) Biology of bats of the New World family Phyllostomatidae, vol Part II. Texas Tech Press, Lubbock, pp. 293–350

Giannini NP, Kalko EKV (2005) The guild structure of animalivorous leaf-nosed bats of Barro Colorado Island, Panama, revisited. Acta Chiropterologica 7(1):131–146. doi:10.3161/1733-5329(2005)7[131:TGSOAL]2.0.CO;2

Giraldeau LA, Valone TJ, Templeton JJ (2002) Potential disadvantages of using socially acquired information. Philos Trans R Soc Lond B Biol Sci 357:1559–1566. doi:10.1098/rstb.2002.1065

Goerlitz HR, Siemers BM (2007) Sensory ecology of prey rustling sounds: acoustical features and their classification by wild grey mouse lemurs. Funct Ecol 21:143–153

Gomes DGE, Page RA, Geipel I, Taylor RC, Ryan MJ, Halfwerk W (2016) Bats perceptually weight prey cues across sensory systems when hunting in noise. Science 353 (6305):1277–1280. doi:10.1126/science.aaf7934

Gridi-Papp M, Rand AS, Ryan MJ (2006) Animal communication: complex call production in the túngara frog. Nature 441:38

Halfwerk W, Dixon MM, Ottens KJ, Taylor RC, Ryan MJ, Page RA, Jones PL (2014a) Risks of multimodal signaling: bat predators attend to dynamic motion in frog sexual displays. J Exp Biol 217(17):3038–3044. doi:10.1242/jeb.107482

Halfwerk W, Jones P, Taylor R, Ryan M, Page R (2014b) Risky ripples allow bats and frogs to eavesdrop on a multisensory sexual display. Science 343(6169):413–416

Handley CO Jr (1976) Mammals of the Smithsonian Venezuelan project. Brigham Young Univ Sci Bull Biol Ser 20:1–89

Håstad O, Victorsson J, Ödeen A (2005) Differences in color vision make passerines less conspicuous in the eyes of their predators. Proc Natl Acad Sci U S A 102(18):6391–6394

Hödl W, Amézquita A (2001) Visual signaling in anuran amphibians. In: Ryan MJ (ed) Anuran communication. Smithsonian Institution Press, Washington, DC, pp. 121–141

Hristov N, Conner WE (2005) Effectiveness of tiger moth (Lepidoptera, Arctiidae) chemical defenses against an insectivorous bat (*Eptesicus fuscus*). Chemoecology 15(2):105–113

Hunt GR, Gray RD (2004) The crafting of hook tools by wild New Caledonian crows. Proc R Soc Lond B Biol Sci 271(Suppl 3):S88–S90. doi:10.1098/rsbl.2003.0085

Ibáñez R, Rand AS, Jaramillo C (1999) The Amphibians of Barro Colorado Nature Monument, Soberania National Park and Adjacent Areas. Mizrachi & Pujol, S.A., Panama

Jones JK Jr (1966) Bats from Guatemala, vol. 16. University of Kansas Publications, Museum of Natural History, pp 439–472

Jones P, Page R, Hartbauer M, Siemers B (2011) Behavioral evidence for eavesdropping on prey song in two Palearctic sibling bat species. Behav Ecol Sociobiol 65(2):333–340. doi:10.1007/s00265-010-1050-9

Jones PL, Farris HE, Ryan MJ, Page RA (2013a) Do frog-eating bats perceptually bind the complex components of frog calls? J Comp Physiol A 199(4):279–283

Jones PL, Ryan MJ, Flores V, Page RA (2013b) When to approach novel prey cues? Social learning strategies in frog-eating bats. Proc R Soc B Biol Sci 280(1772):20132330

Jones PL, Ryan MJ, Page RA (2014) Population and seasonal variation in response to prey calls by an eavesdropping bat. Behav Ecol Sociobiol 68(4):605–615

Jones PL, Ryan MJ, Chittka L (2015) The influence of past experience with flower reward quality on social learning in bumblebees. Anim Behav 101:11–18. doi:10.1016/j.anbehav.2014.12.016

Jones PL, Hämsch F, Kalko EKV, Page RA, O'Mara MT (in review) Foraging and roosting behavior of the fringe-lipped bat, *Trachops cirrhosus*, on Barro Colorado Island, Panama. J Mammal

Kalko EKV, Handley CO, Handley D (1996) Organization, diversity, and long-term dynamics of a Neotropical bat community. In: Cody ML, Smallwood JA (eds) Long-term studies of vertebrate communities. Academic Press, San Diego, pp. 503–553

Kalko EKV, Friemel D, Handley CO, Schnitzler H-U (1999) Roosting and foraging behavior of two Neotropical gleaning bats, *Tonatia silvicola* and *Trachops cirrhosus* (Phyllostomidae). Biotropica 31(2):344–353

Korine C, Kalko EKV (2005) Fruit detection and discrimination by small fruit-eating bats (Phyllostomidae): echolocation call design and olfaction. Behav Ecol Sociobiol 58(5):1–12

Laland KN (2004) Social learning strategies. Learn Behav 32(1):4–14

MacArthur RH, Pianka ER (1966) On optimal use of a patchy environment. Am Nat 100:603–609

Marler PR (1955) Characteristics of some animal calls. Nature 176:6–8

Mettke-Hofmann C (2014) Cognitive ecology: ecological factors, life-styles, and cognition. Wiley Interdiscip Rev Cogn Sci 5(3):345–360

Mikich SB, Bianconi GV, Maia BHLNS, Teixeira SD (2003) Attraction of the fruit-eating bat *Carollia perspicillata* to *Piper gaudichaudianum* essential oil. J Chem Ecol 29(10):2379–2383

Miller GS Jr (1907) The families and genera of bats. Bull US Natl Mus 57:1–282

Morand-Ferron J, Quinn JL (2015) The evolution of cognition in natural populations. Trends Cogn Sci 19(5):235–237

Neuweiler G (1989) Foraging ecology and audition in echolocating bats. Trends Ecol Evol 4:160–166. doi:10.1016/0169-5347(89)90120-1

O'Mara MT, Dechmann DKN, Page RA (2014) Frugivorous bats evaluate the quality of social information when choosing novel foods. Behav Ecol. doi:10.1093/beheco/aru120

Page RA, Ryan MJ (2005) Flexibility in assessment of prey cues: frog-eating bats and frog calls. Proc R Soc B Biol Sci 272:841–847

Page RA, Ryan MJ (2006) Social transmission of novel foraging behavior in bats: frog calls and their referents. Curr Biol 16:1201–1205

Page RA, Ryan MJ (2008) The effect of signal complexity on localization performance in bats that localize frog calls. Anim Behav 76(Part 3):761–769

Page RA, Schnelle T, Kalko EKV, Bunge T, Bernal XE (2012) Reassessment of prey through sequential use of multiple sensory cues by an eavesdropping bat. Naturwissenschaften 99:505–509

Page RA, Ryan MJ, Bernal XE (2014) Be loved, be prey, be eaten. In: Yasukawa K (ed) Animal behavior, Case studies: integration and application of animal behavior, vol 3. Praeger, New York, NY, pp. 123–154

Page RA, Ryan MJ, Kalko EKV, Knörnschild M (in preparation) Sensory mode switching: plasticity in prey detection in the frog-eating bat, *Trachops cirrhosus*

Phillips CJ, Tandler B, Pinkstaff CA (1987) Unique salivary glands in two genera of tropical microchiropteran bats an example of evolutionary convergence in histology and histochemistry. J Mammal 68(2):235–242

Pine RH, Anderson JE (1979) Notes on stomach contents in *Trachops cirrhosus*. Mammalia 43 (4):568–570

Pravosudov VV, Clayton NS (2002) A test of the adaptive specialization hypothesis: population differences in caching, memory, and the hippocampus in black-capped chickadees, *Poecile atricapilla*. Behav Neurosci 116(4):515–522

Ramakers JJC, Dechmann DKN, Page RA, O'Mara MT (2016) Frugivorous bats prefer information from novel social partners. Anim Behav 116:83–87

Rand AS, Ryan MJ (1981) The adaptive significance of a complex vocal repertoire in a Neotropical frog. Zeitschrift fur Tierpsychologie 57(209):214

Ratcliffe JM, ter Hofstede HM (2005) Roosts as information centres: social learning of food preferences in bats. Biol Lett 1:72–74

Rhebergen F, Page RA, Ryan MJ, Taylor R, Halfwerk W (2015) Multimodal cues improve prey localisation under complex environmental conditions. Proc R Soc B Biol Sci 282:20151403

Ripperger S, Josic D, Hierold M, Koelpin A, Weigel R, Hartmann M, Page RA, Mayer F (2016) Automated proximity sensing in small vertebrates: design of miniaturized sensor nodes and first field tests in bats. Ecol Evol. doi:10.1002/ece3.2040

Roberts JA, Taylor PW, Uetz GW (2007) Consequences of complex signaling: predator detection of multimodal cues. Behav Ecol 18(1):236–240

Rodrigues FHG, Reis ML, Braz VS (2004) Food habits of the frog-eating bat, *Trachops cirrhosus*, in Atlantic forest of northeastern Brazil. Chiroptera Neotropical 10:180–182

Rowe C (1999) Receiver psychology and the evolution of multicomponent signals. Anim Behav 58(5):921–931

Russo D, Jones G, Arlettaz R (2007) Echolocation and passive listening by foraging mouse-eared bats *Myotis myotis* and *M. blythii*. J Exp Biol 210:166–176

Ryan MJ, Cummings ME (2013) Perceptual biases and mate choice. Annu Rev Ecol Evol Syst 44:437–459

Ryan MJ, Tuttle MD (1983) The ability of the frog-eating bat to discriminate among novel and potentially poisonous frog species using acoustic cues. Anim Behav 31(3):827–833

Ryan MJ, Tuttle MD, Rand AS (1982) Bat predation and sexual advertisement in a Neotropical anuran. Am Nat 119(1):136–139

Ryan MJ, Tuttle MD, Barclay RMR (1983) Behavioral responses of the frog-eating bat, *Trachops cirrhosus*, to sonic frequencies. J Comp Physiol A Sens Neural Behav Physiol 150(4):413–418

Ryan MJ, Leslie C, Ryan ES (2015) *Physalaemus pustulosus* (túngara frog) sexual communication. Herpetol Rev 46(3):415–416

Schnitzler H-U, Kalko EKV (2001) Echolocation by insect-eating bats. Bioscience 51:557–559

Schnitzler H-U, Moss CF, Denzinger A (2003) From spatial orientation to food acquisition in echolocating bats. Trends Ecol Evol 18(8):386–394

Siemers BM, Schnitzler H-U (2004) Echolocation signals reflect niche differentiation in five sympatric congeneric bat species. Nature 429:657–661

Stuart-Fox D, Moussalli A, Whiting MJ (2008) Predator-specific camouflage in chameleons. Biol Lett 4(4):326–329

Surlykke A, Jakobsen L, Kalko EKV, Page RA (2013) Echolocation intensity and directionality of perching and flying fringe-lipped bats, *Trachops cirrhosus* (Phyllostomidae). Front Physiol 4:143. doi:10.3389/fphys.2013.00143

Tandler B, Phillips CJ, Nagato T (1996) Histological convergent evolution of the accessory submandibular glands in four species of frog-eating bats. Eur J Morphol 34(3):163–168

Taylor R, Ryan M (2013) Interactions of multisensory components perceptually rescue túngara frog mating signals. Science 341(6143):273–274

Taylor RC, Klein BA, Stein J, Ryan MJ (2008) Faux frogs: multimodal signalling and the value of robotics in animal behaviour. Anim Behav 76(Part 3):1089–1097

Toledo RC, Jared C (1995) Cutaneous granular glands and amphibian venoms. Comp Biochem Physiol A Physiol 111(1):1–29

Trillo PA, Athanas KA, Goldhill DH, Hoke KL, Funk WC (2013) The influence of geographic heterogeneity in predation pressure on sexual signal divergence in an Amazonian frog species complex. J Evol Biol 26:216–222

Tuttle MD (2015) The secret lives of bats. Houghton Mifflin Harcourt, Boston, MA

Tuttle MD, Ryan MJ (1981) Bat predation and the evolution of frog vocalizations in the Neotropics. Science 214(4521):677–678

Tuttle MD, Taft LK, Ryan MJ (1982) Evasive behavior of a frog in response to bat predation. Anim Behav 30(2):393–397

Tuttle MD, Ryan MJ, Belwood JJ (1985) Acoustical resource partitioning by two species of Phyllostomid bats (*Trachops cirrhosus* and *Tonatia sylvicola*). Anim Behav 33(4):1369–1371

von Békésy G (1960) Experiments in hearing. McGraw-Hill, New York

Ward AJW, Mehner T (2010) Multimodal mixed messages: the use of multiple cues allows greater accuracy in social recognition and predator detection decisions in the mosquitofish, *Gambusia holbrooki*. Behav Ecol 21(6):1315–1320. doi:10.1093/beheco/arq152

Wright GS, Wilkinson GS, Moss CF (2011) Social learning of a novel foraging task by big brown bats, *Eptesicus fuscus*. Anim Behav 82(5):1075–1083. doi:10.1016/j.anbehav.2011.07.044

Index

M.A. Bee, C.T. Miller (eds.), *Psychological Mechanisms in Animal Communication*, Animal Signals and Communication 5,
DOI 10.1007/978-3-319-48690-1