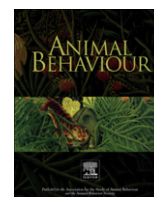


Contents lists available at [SciVerse ScienceDirect](http://www.sciencedirect.com)

Animal Behaviour

journal homepage: www.elsevier.com/locate/anbehav

Essay

Receiver psychology turns 20: is it time for a broader approach?

Cory T. Miller^{a,*,1}, Mark A. Bee^{b,*,1}^a Department of Psychology, University of California San Diego, La Jolla, U.S.A^b Department of Ecology, Evolution and Behavior, University of Minnesota, Twin Cities, St Paul, U.S.A

ARTICLE INFO

Article history:

Received 22 July 2011

Initial acceptance 21 September 2011

Final acceptance 24 October 2011

Available online 19 December 2011

MS. number: AE-11-00584

Keywords:

acoustic communication
 auditory object formation
 auditory scene analysis
 decision making
 receiver psychology
 social categorization
 source segregation
 temporal organization

Twenty years ago, a new conceptual paradigm known as 'receiver psychology' was introduced to explain the evolution of animal communication systems. This paradigm advanced the idea that psychological processes in the receiver's nervous system influence a signal's detectability, discriminability and memorability, and thereby serve as powerful sources of selection shaping signal design. While advancing our understanding of signal diversity, more recent studies make clear that receiver psychology, as a paradigm, has been structured too narrowly and does not incorporate many of the perceptual and cognitive processes of signal reception that operate between sensory transduction and a receiver's response. Consequently, the past two decades of research on receiver psychology have emphasized considerations of signal evolution but failed to ask key questions about the mechanisms of signal reception and their evolution. The primary aim of this essay is to advocate for a broader receiver psychology paradigm that more explicitly includes a research focus on receivers' psychological landscapes. We review recent experimental studies of hearing and sound communication to illustrate how considerations of several general perceptual and cognitive processes will facilitate future research on animal signalling systems. We also emphasize how a rigorous comparative approach to receiver psychology is critical to explicating the full range of perceptual and cognitive processes involved in receiving and responding to signals.

© 2011 The Association for the Study of Animal Behaviour. Published by Elsevier Ltd. All rights reserved.

In the two decades since [Guilford & Dawkins \(1991\)](#) published their now seminal paper on 'receiver psychology', the study of how receivers influence the evolution of animal communication signals has flourished ([Guilford & Dawkins 1993](#); [Rowe 1999](#); [Speed 2000](#); [Candolin 2003](#); [Rowe & Skelhorn 2004](#); [ten Cate & Rowe 2007](#)). A hallmark of the receiver psychology approach has been to elucidate the influence of a receiver's 'psychological landscape' on the evolution of signal design. In introducing the notion of a psychological landscape, [Guilford & Dawkins \(1991, page 2\)](#) described it as 'everything about the brain of the receiver animal that might affect its response to a signal'. Historically, the paradigm emphasizes three features of a receiver's psychological landscape that relate to processes influencing a signal's detectability, discriminability and memorability ([Guilford & Dawkins 1991, 1993](#)). Quite commonly, however, these three features are emphasized only insofar as doing so sheds light on questions about signal evolution. After 20 years, it is fair to ask whether an approach called

'receiver psychology' actually advances our understanding of the 'psychology of receivers', or what [Guilford & Dawkins \(1991, page 2\)](#) rightly referred to as the 'bewilderingly complex system of processors, information-storers and decision-makers' forming the receiver's psychological landscape.

Our objective here is to make the case that the receiver psychology approach is too narrowly conceived and should be broadened to integrate more explicitly the full range of perceptual and cognitive processes that operate from sensory transduction of signals to subsequent behavioural or physiological responses. We aim to do this by highlighting recent research on several processes critical to receiving acoustic communication signals and integrating these topics into a more comprehensive view of receivers' psychological landscapes. As we will illustrate, being on the receiving end of a communicative exchange involves perceptual and cognitive processes unrelated, or at least not closely related, to the ways in which a signal's design determines its detectability, discriminability and memorability. A secondary objective is to emphasize the utility of comparative approaches to the study of receivers' psychological landscapes. Receivers across diverse and distantly related lineages face many of the same fundamental challenges. Consequently, there is potentially wide scope for evolution to produce a diversity of mechanistic solutions to common problems ([Gerhardt & Huber 2002](#)). Explaining signal

* Correspondence: C. T. Miller, Department of Psychology, University of California San Diego, #0109, La Jolla, CA 92093, U.S.A. or M. A. Bee, Department of Ecology, Evolution and Behavior, University of Minnesota, St Paul, MN 55108, U.S.A.

E-mail addresses: corymiller@ucsd.edu (C. T. Miller), mbee@umn.edu (M. A. Bee).

¹ Both authors contributed equally to this work.

diversity has been an important goal of the receiver psychology paradigm for two decades now. Discovering and explaining diversity in the mechanisms for receiving and responding to signals should become a goal of this approach for the coming decades.

A PSYCHOLOGICAL LANDSCAPE FOR ACOUSTIC COMMUNICATION

At a conceptual level, distinctions between 'perceptual' and 'cognitive' processes are ambiguous (Scholl & Leslie 1999), as are those between 'sensory' and 'brain' mechanisms. Sometimes 'sensory system' is misconstrued to mean processing only by the peripheral sense organs. But sensory processing also occurs at much higher levels of the central nervous system (e.g. cortex). Moreover, the modulatory effects of descending efferent innervation on the processing of sensory information at lower levels of the peripheral and central nervous systems are well established (Feng & Ratnam 2000; Møller 2003). It follows that both bottom-up and top-down processes shape many features of a receiver's psychological landscape. Bottom-up mechanisms are often considered to be automatic and obligatory and responsible for processing basic features of direct sensory input. In this sense, they are 'stimulus driven'. In contrast, top-down processing reflects the operation of mechanisms related to an animal's attention, prior experiences and expectations. Top-down processes operate on the output from bottom-up processing, thereby further compounding the difficulty in identifying clear boundaries between perceptual and cognitive components of signal reception. For our discussion, we need not identify such boundaries. Rather, we discuss several processes under the umbrella of perception and cognition that have not been fully integrated into previous conceptualizations of a receiver's psychological landscape, recognizing they may often involve both top-down and bottom-up mechanisms.

We frame our discussion around acoustic communication. The processes we discuss (sound source perception, social categorization, temporal integration and organization, and decision making) occur in the receiver's nervous system between sensory transduction and response output. They represent key aspects of receiving acoustic signals and epitomize the types of perceptual and cognitive processes that deserve consideration in terms of receivers' psychological landscapes. Investigating these features has been fundamental in research on human hearing, speech communication and cognition, and we draw heavily on this literature. An important point to bear in mind is that many of the issues we touch upon are not unique to receiving acoustic signals and have parallels in other sensory modalities, in particular, vision (Miller & Cohen 2010). We purposefully focus on acoustic communication systems, however, because they are among the most comprehensively studied (Gerhardt & Huber 2002; Simmons et al. 2003; Catchpole & Slater 2008; Bradbury & Vehrencamp 2011). In addition, the potential for diversity in receiver solutions to common problems seems high in this modality. The sense of hearing evolved numerous times (Webster et al. 1992), and even within lineages some key features of audition (e.g. tympanic hearing among tetrapod vertebrates) had multiple evolutionary origins (Christensen-Dalsgaard & Carr 2008). Thus, even if receivers across divergent taxa appear to behave similarly in similar situations, there can be no guarantee that precisely the same underlying auditory mechanisms are at work.

In behavioural studies of communication, our only windows into a receiver's psychological landscape are its behavioural or physiological responses (or lack thereof) to signals. Thus, a key challenge to implementing experimental research on perceptual and cognitive processes in acoustic communication systems is isolating them in the context of complex and often reciprocal

behavioural interactions between signallers and receivers. The studies described below did so successfully using techniques ranging from field playback experiments eliciting natural behaviours in natural habitats to traditional psychoacoustic methods based on conditioned responses in laboratory sound chambers.

SOUND SOURCE PERCEPTION

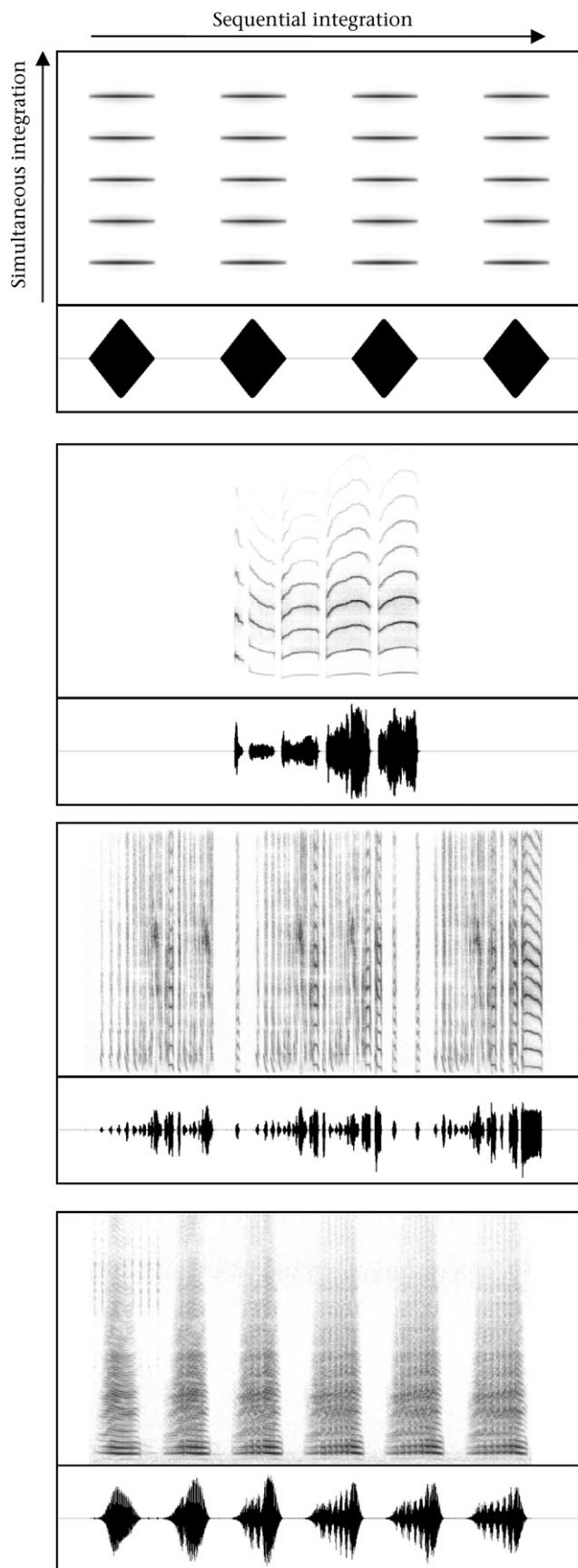
A major focus of research on human hearing and speech communication has aimed to understand how we experience sounds as distinct sources in the environment (reviewed in Yost et al. 2008). This is not a trivial matter for the auditory system. Sound pressure waves generated by multiple concurrent sources in the environment add together to form a composite pressure wave that impinges upon the ears. Audition is essentially tasked with decomposing this composite waveform and assigning its constituent parts to coherent perceptual representations of the different sound sources active in the environment. Sound source perception is thus critical for communicating acoustically with conspecifics, as well as for navigating a world full of other potentially beneficial (prey) or harmful (predator) sound sources. By comparison with humans, we know relatively little about sound source perception in nonhuman animals (Hulse 2002; Bee & Michéyl 2008). Two critical aspects of sound source perception discussed in this section involve the *integration* of sounds arising from the same source or event into perceptually distinct units ('objects') and their *segregation* from other sounds in the environment, especially the background noise generated by other signalling animals.

Signals as Perceptual Objects

The basic unit of perception is the 'object' (Spelke et al. 1993; Scholl 2001). Based on work primarily in the visual system, objects are loosely considered spatio-temporally bounded feature clusters (Scholl 2001). For something to be perceived as an object, it must move through space and time as a single coherent, bounded unit. Forming perceptual objects involves binding features *x*, *y* and *z* into a coherent representation of 'object A'; this is not the same as detecting that 'object A' is present (versus absent) or discriminating between 'object A' and 'object B' based on differences in the values of features *x*, *y* or *z*. Object perception refers to mechanisms by which nervous systems bind those features together that belong together. The concept of an 'auditory object' has been used to describe coherent representations that correspond both to sound sources (e.g. a voice) and to specific acoustic events (e.g. a vocalization) (Griffiths & Warren 2004; Miller & Cohen 2010). While forming auditory objects seems critical to receiving acoustic communication signals (Darwin 2008), we know little about how nonhuman receivers accomplish this task. An expanded view of receiver psychology considers communication signals or their sources as perceptual objects in a receiver's environment and the processes leading to their formation (e.g. Gentner 2008; Miller & Cohen 2010).

Auditory object formation

Two general forms of perceptual grouping in auditory object formation can be distinguished (Bregman 1990). 'Sequential integration' involves the grouping of temporally separated sounds from the same source (e.g. notes, song motifs, words) into a coherent 'auditory stream' that unfolds through time and can be selectively attended to by receivers (Fig. 1). 'Simultaneous integration' refers to the perceptual binding of different, simultaneously occurring spectral components from one source (e.g. harmonics, formants) into a single coherent representation. Both forms of grouping can be illustrated with a musical example. After hearing a pianist play



'middle C' (C_4), we perceive a single, integrated note having a certain pitch and not separate, simultaneous notes corresponding to the multiple harmonics of a fundamental frequency of 261.6 Hz. This would be an example of simultaneous integration. Our ability to recognize temporally separated notes played on a piano as a coherent melody unfolding through time represents sequential integration in action. As illustrated in Fig. 1, both sequential and simultaneous integration may be critical for the reception of acoustic signals in many animals (Hulse 2002; Bee & Micheyl 2008).

Both bottom-up and top-down processes contribute to forming auditory objects in humans (Bregman 1990; Feng & Ratnam 2000; Näätänen et al. 2001; Carlyon 2004). Bottom-up mechanisms appear to analyse a relatively small number of stimulus features that function as low-level cues for grouping sound elements arising from the same source (see Fig. 2). These grouping cues include common onsets and offsets, harmonicity (i.e. common fundamental frequency, F_0), common patterns of amplitude modulation, spectral and temporal proximity, common timbre, and perhaps to a lesser extent, common spatial location (reviews in Hartmann 1988; Bregman 1990; Darwin & Carlyon 1995; Darwin 1997; Bee & Micheyl 2008). The interplay between these stimulus-driven cues and the modulating effects of top-down processing are nicely illustrated in studies of 'auditory streaming' in humans using simple tones to investigate sequential integration (Bregman 1990; Carlyon 2004; Snyder & Alain 2007).

When presented with rapid sequences of two alternating tones (A and B) arranged as repeating triplets (ABA—, where the dash indicates silence), we can have two distinct perceptual experiences depending on the frequency separation (ΔF) between the A and B tones (van Noorden 1975). When the two tones are similar in frequency (i.e. high spectral proximity), we tend to hear a single, integrated sequence with a galloping rhythm comprising tones jumping up and down in frequency (e.g. ABA—ABA—ABA—...; Fig. 3a). As spectral proximity decreases, however, our percept changes dramatically. At sufficiently large frequency separations, the A and B tones perceptually 'split' into separate 'auditory streams' consisting of only A tones or only B tones, each with an isochronous rhythm, one half the rate of the other; the galloping rhythm is lost (e.g. A—A—A—A—A—A—... and —B—B—B—B—B—B—...; Fig. 3b). The tendency to hear two streams also generally increases with faster tone sequences as a result of increased temporal proximity between consecutive tones of the same frequency.

Three important points are worth making with this example. First, the minimum frequency separation causing the tones to split into separate streams depends on the instructions given to listeners (Fig. 3c). When asked to listen analytically for the isochronous rhythms corresponding to the split, two-stream percept, subjects report hearing two streams at much smaller frequency separations

Figure 1. Spectrograms (1024 points, 20–48 kHz sampling rates) illustrating the concepts of sequential and simultaneous integration in relation to the vocalizations of several animal species. Many animals, such as those depicted here, produce temporally structured acoustic signals with distinct sound elements (e.g. pulses, notes, motifs) separated from each other by brief periods of silence. In addition, each sound element often contains multiple concurrent spectral components (e.g. harmonics or formants). As indicated by the arrows, sequential integration refers to the perceptual binding of temporally separated sound elements across time and simultaneous integration refers to the binding of concurrent elements produced by the same source. From top to bottom are shown a repeated complex tone with five harmonics ($F_0 = 1$ kHz), the combination long call of a cottontop tamarin, *Saguinus oedipus*, the song of a zebra finch, *Taeniopygia guttata*, and the advertisement call of a North American bullfrog, *Lithobates catesbeianus*. In all spectrograms, the X axis depicts 8 s of sound; from top to bottom, the Y axes depict frequency ranges extending up to 6 kHz, 20 kHz, 10.5 kHz and 5 kHz, respectively. All depicted sounds were generated or recorded by the authors except for the zebra finch song (courtesy Teresa Nick).

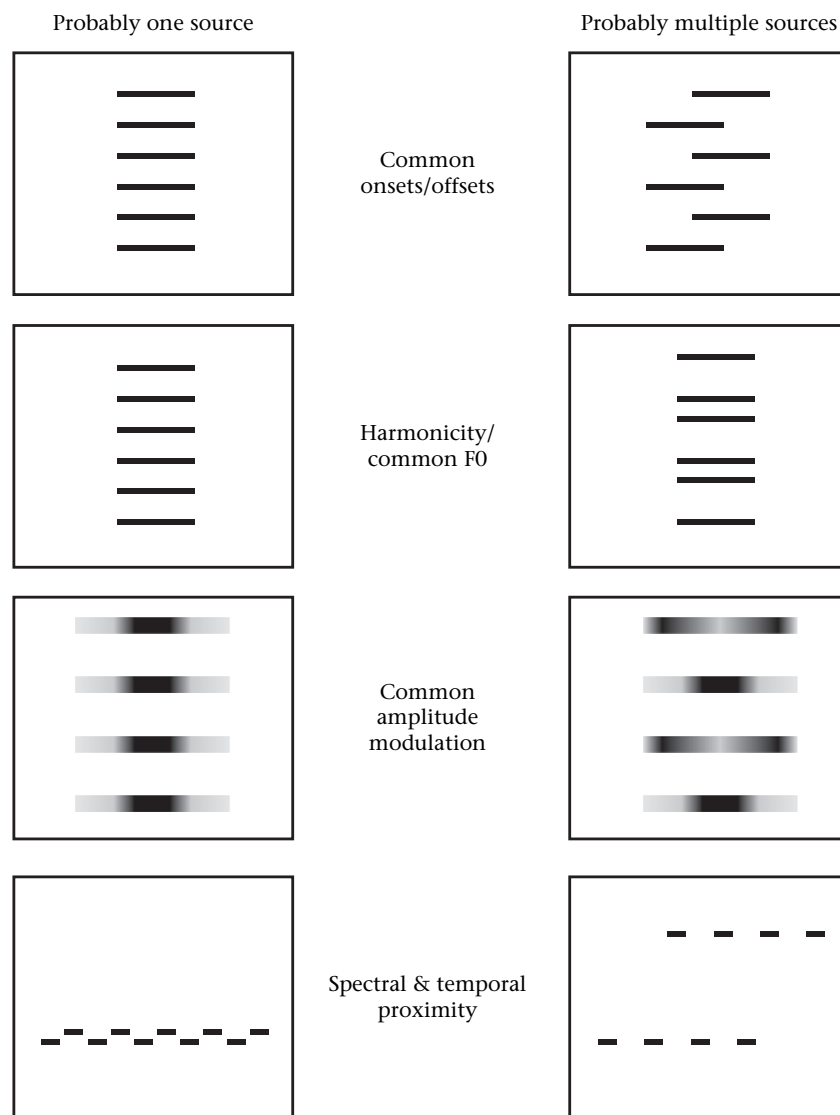


Figure 2. Stylized spectrograms plotting frequency (Y axis) versus time (X axis), with amplitude reflected in the darkness of coloration (Z axis) and illustrating several potent acoustic cues for auditory grouping.

(the ‘fission boundary’ in Fig. 3c) compared to when they are asked to listen instead for the integrated, galloping rhythm (the ‘temporal coherence boundary’ in Fig. 3c). Second, at some intermediate frequency separations, listeners can wilfully shift their attention between the two percepts, an acoustic example of perceptual bistability (Fig. 3c). Together, these two observations illustrate the interplay between bottom–up and top–down mechanisms. Bottom–up mechanisms related to frequency selectivity, forward suppression and neural adaptation process stimulus properties at low levels of the auditory system and contribute to forming two-stream percepts by segregating neural responses to tones of different frequency in tonotopic space (Fishman et al. 2001, 2004; Bee & Klump 2004, 2005; Pressnitzer et al. 2008; Bee et al. 2010). Yet in some cases, top–down processes related to the listener’s attention and expectations influence what is ultimately perceived (i.e. one or two streams) by determining whether the A and B tones are assigned to the same or different auditory streams. Finally, whether listeners hear one or two auditory streams is not a simple matter of detecting the presence or absence of A versus B tones, or discriminating between them; two suprathreshold tones differing

in frequency may be highly discriminable yet fail to segregate into separate streams (Rose & Moore 2005).

Forming perceptual objects of signals

Assigning the interleaved or overlapping signal elements produced by two or more nearby animals to different auditory objects would seem fundamental for many acoustically mediated behaviours, such as finding relatives in a colony (Aubin & Jouventin 2002) or eavesdropping on signalling interactions in a communication network (Langemann & Klump 2005; Peake 2005). Differences in pitch related to differences in fundamental frequency (i.e. reduced spectral proximity) allow human listeners to segregate not only interleaved tone sequences (Fig. 3), but also concurrent, overlapping voices (Brokx & Nootboom 1982; Bird & Darwin 1998). Studies using simplified tone-like sounds in psychophysical studies of goldfish (Fay 1998), starlings (MacDougall-Shackleton et al. 1998), ferrets (Ma et al. 2010) and monkeys (Izumi 2002) have shown these nonhuman animals also possess capabilities of auditory streaming. Only two studies, however, have investigated auditory streaming in the context of nonhuman

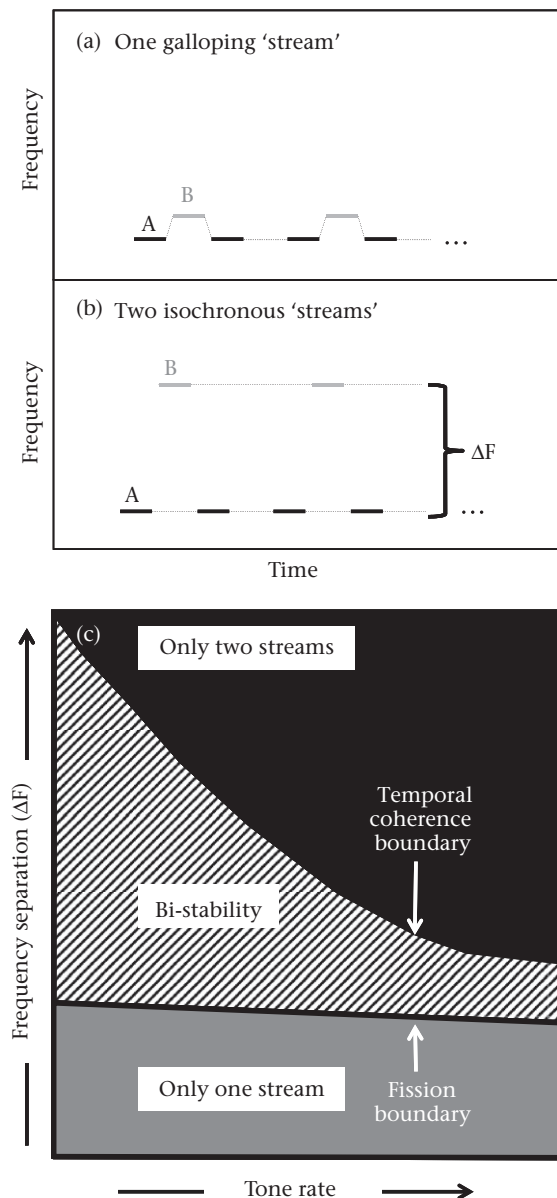


Figure 3. Schematic illustration of auditory streaming and the interplay between top-down and bottom-up influences. (a) When two tones (A and B) are arranged as triplets with small frequency differences (ΔF), we hear a distinctive galloping rhythm. (b) When ΔF is increased, the A and B tones are perceptually segregated into separate auditory streams, each with its own steady, isochronous rhythm. (c) The effects of ΔF (and also tone rate) on auditory streaming depend on how a listener directs their attention. When subjects are asked to listen for the isochronous rhythms, they can do so at the ΔF s indicated along the fission boundary. Below the fission boundary, it is not possible to hear two segregated streams (that is, the two streams are 'fused' into one stream). However, when instructed to listen instead for the galloping rhythm, they can do so up to the higher ΔF s along the temporal coherence boundary. Between the fission and temporal coherence boundaries, the one-stream and two-streams percepts exhibit bistability and listeners can voluntarily shift their attention between the two percepts. The location of the fission boundary is thought to be largely determined by relatively low-level, stimulus-driven mechanisms (e.g. frequency selectivity and forward suppression; Bee & Klump 2004). The difference in location between the fission and temporal coherence boundaries reflects the influence of top-down mechanisms. The schematic representation of fission and temporal coherence boundaries in (c) after van Noorden (1975); the X axis depicts a range of about 100 ms and the Y axis depicts a range of about one octave.

animal communication. Schul & Sheridan (2006) were first to do so in showing that auditory interneurons in the katydid, *Neoconocephalus retusus*, can segregate streams of conspecific songs and bat echolocation calls based on differences in their spectrotemporal structures. Nityananda & Bee (2011) recently showed that auditory stream segregation based on naturally occurring frequency differences facilitates receiving conspecific signals amid the many overlapping calls in mixed-species frog choruses.

In addition to spectral proximity, there is accumulating evidence that other stimulus-driven grouping cues exploited by humans also operate in the communication systems of nonhuman animals. For example, Geissler & Ehret (2002) showed that common onsets of harmonics were an important component of the structure of mouse (*Mus domesticus*) pup wriggling calls that promoted simultaneous integration by receivers. Shifting the timing of the first harmonic by only 20–30 ms before or after other harmonics reduced the signal's ability to elicit stereotyped maternal behaviours. Dent et al. (2009) recently showed that common onsets may have much stronger influences in promoting auditory grouping in birds than in humans tested under similar listening conditions. Harmonicity (i.e. common F0) appears to influence call perception in cottontop tamarins, *Saguinus oedipus* (Weiss & Hauser 2002), and North American bullfrogs, *Lithobates catesbeianus* (Simmons & Bean 2000). Studies of several other frog species suggest disrupting common spatial origin by presenting different components of the same signal from different locations can have large impacts on simultaneous integration (Bee 2010), but perhaps somewhat weaker influences on sequential integration (Schwartz & Gerhardt 1995; Gerhardt et al. 2000; Farris et al. 2002, 2005; Bee & Riemersma 2008). The effects of spatial origin on sequential integration are more pronounced when frogs are forced to choose which of two simultaneous elements to group with a preceding sound (Farris & Ryan 2011).

Filling in the gaps

Sometimes receivers may encounter situations in which they lack all of the acoustic information needed to form coherent auditory objects, such as when part of a signal is masked by a loud but sudden and brief sound. One example familiar to scientists might be when a nearby audience member at a scientific meeting coughs at a key moment during a plenary talk. Studies of 'phonemic restoration' (Warren 1970) in humans have shown we are quite adept at filling in gaps in partially masked speech created by such intermittent bursts of noise (reviewed in King 2007). In fact, human listeners can perceive correct but illusory phonemes in speech tokens in which the phoneme has actually been deleted and replaced with masking noise. More broadly termed 'auditory induction' (reviewed in King 2007), similar perceptual illusions have also been demonstrated using acoustic communication signals in monkeys (Miller et al. 2001; Petkov et al. 2003) and songbirds (Braaten & Leary 1999; Seeba & Klump 2009). Hence, receivers in these taxa also possess mechanisms for reconstructing percepts of signals in the face of incomplete or degraded acoustic information. In humans, prior experience with specific words (i.e. a top-down influence) can affect the strength of perceptual restoration (Samuel 1996). At least one study of songbirds suggests top-down influences related to prior experience may also influence the strength of perceptual restoration in some nonhuman animals (Seeba & Klump 2009). In contrast to these earlier studies of humans, monkeys and songbirds, studies of frogs have so far failed to reveal strong evidence for auditory induction (Schwartz et al. 2010; Seeba et al. 2010). These taxonomic differences highlight the potential for diversity in receivers' psychological landscapes.

Exploiting Noise to Process Signals

For many animals, the perceptual analysis of acoustic scenes requires segregating one or more behaviourally relevant signals from high levels of background noise in social aggregations or communication networks (Hulse 2002; Brumm & Slabbekoorn 2005; Langemann & Klump 2005; Bee & Micheyl 2008). In noisy social environments, the acoustic signals of other conspecifics represent particularly potent sources of both energetic and informational masking (reviewed in Brumm & Slabbekoorn 2005). The difficulty we experience in understanding speech by one person in multi-talker groups has been termed the 'cocktail-party problem' (Cherry 1953; reviewed in McDermott 2009). As we illustrate in this section, detecting and discriminating among signals in such environments may be influenced as much by features of the noise and its relationship to signals as by the design features of signals themselves. That is, detectability and discriminability are not always simple matters of optimal design to maximize signal efficacy. Instead, receivers' psychological landscapes include mechanisms that exploit certain spatial and temporal relationships between signals and noise to effect a 'release' from auditory masking.

Spatial release from masking

Separating target speech from speech-like noise by 90° in azimuth substantially reduces speech recognition thresholds (e.g. by 6–10 dB; reviewed in Bronkhorst 2000). Known as 'spatial release from masking', our ability to exploit spatial separation between signals and noise facilitates speech comprehension in noisy social settings. Both bottom-up mechanisms (e.g. processing stimulus-driven binaural cues) and top-down mechanisms (e.g. spatial attention) have been implicated in this process in humans (Shinn-Cunningham et al. 2005). Studies of frogs (Schwartz & Gerhardt 1989; Bee 2007, 2008; Richardson & Lengagne 2010; Nityananda & Bee, in press) and birds (Dent et al. 2009) suggest spatial release from masking might also contribute to functionally important behaviours that depend on segregating individual calls from the background noise of an aggregation. For example, Dent et al. (2009) trained zebra finches, *Taeniopygia guttata*, and budgerigars, *Melopsittacus undulatus*, which commonly occur in mixed-species flocks, to categorize the songs of different individual zebra finches. In the presence of biologically realistic noise, subjects reached threshold levels of correct performance at lower (e.g. by 20–30 dB) target-to-masker ratios when signals and noise were separated by 90° compared to a co-localized condition. In studies of grey treefrogs, *Hyla chrysoscelis*, females experienced masking release (about 3–6 dB) when a target signal and a source of chorus-like noise were separated by 90° compared to co-localized conditions (Bee 2007; Nityananda & Bee, in press). Spatial unmasking also improved females' discrimination between conspecific and heterospecific calls in chorus noise (Bee 2008).

Dip listening

Receivers can also experience masking release by exploiting temporal fluctuations in the level of background noise. Human listeners, for example, are better able to recognize speech in the presence of speech-like noise when the masker fluctuates in level over time, thereby making speech more audible during brief 'dips' in amplitude (reviewed in Füllgrabe et al. 2006). Studies of this so-called 'dip listening' in the nightingale grasshopper, *Chorthippus biguttulus* (Ronacher & Hoffmann 2003), and grey treefrogs, *H. chrysoscelis* (Vélez & Bee 2011), have shown that both invertebrates and nonhuman vertebrates also experience masking release in temporally fluctuating noise. We might expect masking release in temporally structured noise to be prevalent among receivers,

because temporal fluctuations are a prominent feature of background noise (Klump 1996; Nelken et al. 1999; Vélez & Bee 2010).

Two additional points about temporal fluctuations in background noise are pertinent here. First, different parts of the frequency spectrum of environmental noises may undergo similar fluctuations in level over time; that is, temporal fluctuations in the amplitude of noise may often be correlated across frequencies (i.e. 'comodulated'; Klump 1996; Nelken et al. 1999). Humans can exploit comodulation in background noise to detect signals at lower levels (reviewed in Verhey et al. 2003). Although not yet investigated in nonhuman animals using communication signals, evidence from behavioural studies of fish (Fay 2011), birds (Klump & Langemann 1995; Langemann & Klump 2001, 2007; Jensen 2007) and dolphins (Branstetter & Finneran 2008) suggest receivers may be able to exploit comodulation in acoustic signal perception (for an example with speech perception in humans, see Kwon 2002). Second, the sounds generated by large aggregations of signalling animals can also exhibit species-specific differences in the temporal patterning of amplitude fluctuations that reflect underlying differences in both signal structure and the calling behaviour of individuals composing the aggregation (Vélez & Bee 2010). The extent to which receivers possess mechanisms specialized for exploiting these species-specific level fluctuations (as opposed to more generalized dip-listening abilities) remains an open question.

Multimodal Effects in Sound Source Perception

Increased study of multimodal signalling has been one of the major contributions of the receiver psychology approach (Rowe 1999; Rowe & Guilford 1999). Perhaps not surprisingly, perceiving sound sources is subject to multimodal influences. Among the best known examples of such multimodal interactions in humans include our reliance on lip reading to understand speech under noisy listening conditions (Sumbly & Pollack 1954), the McGurk effect demonstrating the combined influence of acoustic and visual information on phoneme determination (McGurk & Macdonald 1976), and the remarkable success of professional ventriloquists in tricking us to perceive talking puppets despite some spatial separation between acoustic and visual components of speech (Alais & Burr 2004). Nonhuman animals also perceptually integrate acoustic signals with visual cues associated with their production, such as an inflating vocal sac in frogs (Rosenthal et al. 2004; Narins et al. 2005; Taylor et al. 2008, 2011) and appropriate facial movements in macaque monkeys (Ghazanfar & Logothetis 2003) and chimpanzees (Izumi & Kojima 2004). These studies serve to illustrate that mechanisms for cross-modal integration in sound source perception may be important features of a receiver's psychological landscape across diverse taxa. Describing the actual physiological mechanisms of cross-modal sensory integration across divergent species will no doubt be a fruitful avenue for future investigations into multimodal signalling (Romanski & Ghazanfar 2009).

Summary

The handful of available studies on sound source perception in insects, frogs, birds and mammals highlight the potential for diversity in receiver mechanisms. This diversity, in turn, could provide a basis for some interesting comparative investigations into similarities and differences in how animals receive signals. Although acoustic signals are structured according to the physical constraints governing sound production, there is potential scope for evolutionary differences in receiver psychology related to how receivers construct meaningful percepts of those signals amid so many other competing signals and high levels of ambient noise. Describing this diversity in mechanistic terms is necessary to

understand the evolution of communication systems. In addition, we are unaware of any studies of nonhuman animals explicitly designed to distinguish between the operation of top–down and bottom–up processes in the joint context of sound source perception and communication. There may be important differences awaiting discovery concerning the roles of bottom–up and top–down processes in sound source perception across taxa. We would predict top–down influences to play particularly important roles among vocal learners and in animals exhibiting experience-based social categorization (next section). Another important goal for future research under the receiver psychology paradigm should be to determine the extent to which mechanisms for sound source perception might both drive and constrain signal design and signalling behaviours.

SOCIAL CATEGORIZATION

Receivers must often categorize signallers along one or more dimensions based on features of their signals as well as the social context in which signals are received. This key function of communication has obvious ties to a signal's discriminability and memorability. However, describing signal-based social categorization simply as a function of the ability of receivers to discriminate and remember signals glosses over some potentially interesting features of, and species differences in, receivers' psychological landscapes.

A General Framework

The foundations of social categorization are likely based on more generalized cognitive capacities for categorization. Following Rosch (Rosch et al. 1976; Rosch 1978), objects can be categorized at multiple levels along a general-to-specific hierarchy: superordinate, basic-level and subordinate (Fig. 4). A visual object, such as a coffee mug, can simultaneously belong to a superordinate category of 'container', a basic-level category of 'mug', and subordinate categories of 'holds hot liquid' and 'the colour white'. This hierarchy for categorization borrowed from human cognitive psychology provides a common framework for comparing how receivers in different taxonomic groups accomplish the general task of categorizing other individuals by their signals. For a young male songbird being tutored by his father, for example, a particular vocalization could concurrently belong to the superordinate category of 'conspecific song', the basic-level category of 'song type A' and subordinate categories of 'male song' and 'parent's song' (Fig. 4).

Assignments of signals to superordinate categories (e.g. conspecific versus heterospecific) and basic-level categories

(e.g. signal types) may be driven primarily by the recognition of spectrotemporal feature combinations idiosyncratic to the species and signal type (e.g. Nelson 1988; Searcy et al. 1999; Gerhardt 2001; Weiss & Hauser 2002; Miller et al. 2004, 2005; Baugh et al. 2008). For some animals (most likely birds and mammals), learning and experience may contribute significantly to categorization at these levels; for many others, however (e.g. insects, fish, frogs), assigning signals to superordinate and basic-level categories may require no previous exposure to signals at all, and may be accomplished using ever more-selective peripheral and central filters tuned to specific acoustic features. Identifying mechanisms for species and call type recognition has been a hallmark of neuroethological studies of animal communication (Gerhardt & Huber 2002).

A central component of social categorization is the ability of receivers to categorize conspecifics at one or more subordinate levels, such those related to a caller's sex, size or dialect (Gerhardt 1992). Among the best-studied subordinate level categories is individual identity.

Categorizing Individuals

Acoustic signals are individually distinct, and animals as diverse as fish (Myrberg & Riggio 1985), frogs (Bee & Gerhardt 2001a), seabirds (Aubin & Jouventin 2002), songbirds (Brooks & Falls 1975; Nelson 1989), bats (Balcombe 1990), monkeys (Rendall et al. 1996) and dolphins (Sayigh et al. 1999) recognize individual conspecifics based on patterns of individual variation in acoustic signals (reviewed in Bee 2006). While we know that many animals recognize familiar individuals by their acoustic signals, we know far less about the underlying psychological mechanisms of individual recognition.

In some animals, for example, subordinate categorization at the level of specific individuals is based on the ability of receivers to learn idiosyncratic features of a signaller's voice, such as those related to pitch. This appears to be the case for some songbirds (e.g. Brooks & Falls 1975; Nelson 1989) and frogs (e.g. Bee & Gerhardt 2001a, 2002). But 'voice recognition' per se is not the only mechanism enabling subordinate categorization of specific individuals. In European starlings, *Sturnus vulgaris*, receivers can recognize individuals by memorizing large sets of individually distinctive song types and their sequential ordering within song bouts (Gentner & Hulse 1998, 2000; Gentner 2004). Recognition, therefore, is based on the receiver's ability to organize the song's overall temporal structure into a hierarchy of individually distinctive bout sequences (Gentner 2008; Comins & Gentner 2010). Because of this hierarchical structure, starlings need only to hear a short sequence of motifs in order to recognize an individual's identity (Knudsen et al. 2010). The available work on frogs and songbirds suggest that a diversity of cues, and hence mechanisms for processing them, may be involved in subordinate categorization of specific individuals.

An exclusive focus on signals and their discriminability and memorability misses some other important aspects of social categorization by receivers. For example, recent work with domestic horses, *Equus caballus*, and rhesus macaques, *Macaca mulatta*, indicates that individual recognition in these species is cross-modal, meaning that receivers have a complex, multidimensional representation of known individuals that includes not only information on idiosyncratic features of their vocalizations, but also what they look and/or smell like (Proops et al. 2009; Sliwa et al. 2011). This is perhaps quite similar to our own human experience of individual recognition, which is also multidimensional. As these studies illustrate, social categorization at the level of the individual is not necessarily a simple function of a signal's discriminability and memorability within a particular sensory modality.

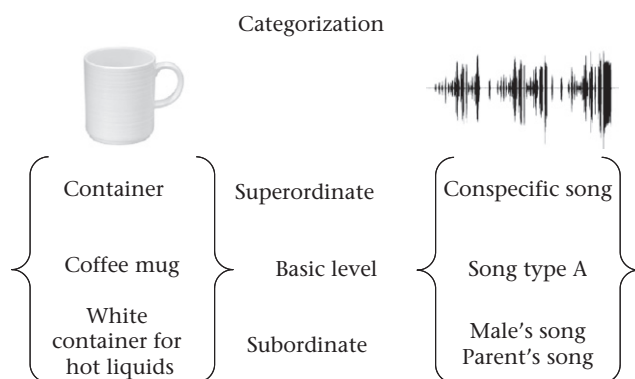


Figure 4. Three levels of categorization following Rosch (1978) and illustrated for a coffee mug and a zebra finch song.

Categorizing Social Relationships

In addition to multidimensional representations of specific individuals, receivers in some species also may possess rich social knowledge about individuals that can be accessed upon receiving a signal or hearing signals in particular contexts. Experimentally manipulating the context of signal reception may often be the only way to probe the boundaries of receivers' social categories and knowledge about relationships among categories. An elegant set of field playback experiments conducted on chacma baboons, *Papio hamadryas ursinus*, illustrates the significance of context to social categorization. In this species, dominant females typically produce a grunt when approaching a subordinate female. A common response of the subordinate female in this situation is to produce a fear bark. In a field playback experiment, [Cheney et al. \(1995\)](#) presented a sequence of calls that was socially consistent (grunt of a dominant followed by the fear bark of a subordinate), while in a second condition they broadcast a socially inconsistent sequence of calls (grunt of a subordinate followed by the fear bark of a more dominant individual). Subjects were significantly more responsive during the socially inconsistent condition, suggesting they not only recognized the two individuals, but that vocal recognition also led to recall of knowledge about their relative social rank. [Bergman et al. \(2003\)](#) subsequently showed that baboons categorized not only the relative rank of all individuals across the group, but their relative position within and between their respective matrilineas as well. Together these data suggest that baboons recognize not just individuals but also their place in more complex social networks. Importantly, these social categorizations are inextricably bound to the perception of acoustic signals in particular contexts.

Summary

Mechanisms for social categorization represent important features of receivers' psychological landscapes. Broadening our conception of receiver psychology would provide a valuable framework for considering the multiple levels at which signals can be categorized in social recognition, the neural mechanisms by which these categorizations are made, and the social knowledge receivers can access in memory based on receiving those signals. Given the prevalence of individual vocal recognition in animals, comparative studies of social categorization at the subordinate level of individuals might bring to light some interesting similarities and differences in categorization across diverse taxa. By taking a broad taxonomic approach, it may be possible to evaluate the relationship between the sophistication of social categorization and the requisite underlying mechanisms for this aspect of receiver psychology.

TEMPORAL INTEGRATION AND ORGANIZATION

Another important aspect of a receiver's psychological landscape, and one potentially tied more directly to detectability, discriminability and memorability, involves the integration and organization of acoustic information over behaviourally relevant timescales. The basic act of detecting acoustic signals requires integration of sensory input over time ([Heil & Neubauer 2003](#)). But beyond hearing threshold determinations, receivers' abilities to integrate and organize acoustic information over different timescales are critical for communication in many different contexts. For example, Pacific treefrogs, *Pseudacris regilla*, integrate a minimum number of pulses produced at the correct rates to make basic-level categorizations of advertisement calls and encounter calls ([Rose & Brenowitz 2002](#)), a process that likely involves midbrain neurons that count ([Alder & Rose 1998](#); [Edwards et al. 2002](#)). Emperor penguins, *Aptenodytes*

forsteri, and cottontop tamarins (*S. oedipus*) must integrate information over minimum numbers of call syllables before they can categorize conspecifics to the subordinate level of specific individuals ([Aubin & Jouventin 2002](#); [Miller et al. 2005](#)). Many animals modify the temporal sequence of signal elements (e.g. syllables or call types) in various contexts (e.g. [Robinson 1979](#); [Searcy 1983](#); [Payne & Payne 1985](#); [Mitani & Marler 1989](#); [Searcy et al. 1995](#)). Evidence that these signal modifications subsequently alter what is actually communicated, and thereby function as syntax, is more limited ([Searcy & Nowicki 1999](#); [Freeberg & Lucas 2002](#); [Zuberbühler 2002](#); [Clucas et al. 2004](#)). Nevertheless, syntax would require receivers to organize signal information over timescales longer than individual signal elements.

Working Memory and Change Detection

A primary cognitive constraint to consider is the temporal window over which sounds can be organized into meaningful units of communication. Fundamental psychological processes, such as change/novelty detection ([Simons 1996](#); [Ranganath & Rainer 2003](#)) and working memory ([Baddeley 2003](#)), may ultimately be the foundation for the mechanisms underlying the integration of temporal information during acoustic interactions. For example, a recent study of European starlings (*S. vulgaris*) found that this species' ability to use sequences of motifs for individual recognition is made possible, at least in part, by holding series of acoustic units in working memory ([Comins & Gentner 2010](#)). Other work with this same species has identified differences in the persistence times of auditory memories depending on whether the birds were required to remember spectral or temporal acoustic information ([Zokoll et al. 2007, 2008](#)). Working memory also influences the mating decisions that female túngara frogs, *Engystomops pustulosus*, make in response to males that add multiple 'chucks' to the simple 'whine' component of the species-specific advertisement call ([Akre & Ryan 2010](#)). Such processes may be involved in many aspects of acoustic communication, such as the stimulus-specific habituation hypothesized to underlie vocally mediated dear enemy recognition in some territorial animals ([Bee & Gerhardt 2001b](#)).

Tonic Signalling

The emission of a continuous sequence of signals or reciprocal signal exchanges over relatively long periods of time (e.g. minutes or hours), known as 'tonic signalling' ([Schleidt 1973](#)), represents an example of acoustic communication that requires receivers to integrate information over temporal windows much longer than individual signals. Tonic signalling can convey information both in a monotonous sequence of signals and in changes in these sequences, such as alterations in the temporal pattern or even the complete cessation of signalling. In Richardson's ground squirrels, *Spermophilus richardsonii*, for example, evenly spaced elements in tonic alarm calling increase vigilance in groupmates, while a more variable temporal patterning leads to a decrease in the same behaviours ([Sloan & Hare 2004](#)). Tonic production of contact calls in bobwhite quail, *Colinus virginianus*, seems to communicate a lack of danger as group members are foraging in dense brush, while a cessation of calling communicates the presence of danger, leading individuals to freeze ([Stokes 1967](#)). In túngara frogs (*E. pustulosus*), the cessation of repeated calling by one or a few males simulated by playbacks was sufficient to silence small choruses ([Dapper et al. 2011](#)). In highly voluble species with diverse acoustic repertoires, such as many New World primates, information may commonly be conveyed not only by individual vocalizations, but also in the pattern of tonic calling across individuals in the group. The significance of tonic signalling in animal communication systems is not

fully known, in part, because emphasis is more typically placed on what might be called 'phasic signalling' and direct functional relationships between individual signals and a receiver's immediate behavioural response. Considerably more work is needed on tonic communication systems, as they may reveal particularly dynamic cognitive features of a receiver's psychological landscape in some species.

Summary

Organizing acoustic signals along a temporal dimension is central to many elements of communication behaviour. For acoustic communication, signalling and receiving can never be divorced from considerations of time. What is striking to us is that time is inextricably connected to so many different aspects of signal reception and a receiver's psychological landscape, from influencing signal detection thresholds to forming auditory objects, discriminating between call types and recognizing specific individuals. We believe characterizing these diverse processes in terms relating only to signal detectability, discriminability and memorability potentially misses many interesting questions related to diversity in the underlying mechanisms for organizing and integrating signals over time.

DECISION MAKING

It is often convenient, and has been common historically, to conceive of 'acts' of communication as discrete exchanges of information between one signaller and one receiver with well-defined beginnings (a signaller signals) and outcomes (a receiver responds). But communication is frequently interactive and dynamic, and social environments for signalling are typically quite complex (McGregor 2005). In light of these circumstances, decision making is an important aspect of a receiver's psychological landscape. While animals must 'decide' whether a signal is present, whether two signals differ, or whether a signal matches a learned template, many other decisions seem to go beyond simple issues of detectability, discriminability and memorability.

Whether, What and When?

Upon receiving signals, animals face decisions about whether, what and when behaviour will be generated in response (Lee 2010). These decisions differ in their range of possible options. The decision about *whether* to respond is binary (i.e. a response is generated or it is not). The occurrence of a response can be influenced by a suite of ecological and social factors, such as whether the individual perceives the initial signal as being directed towards them or another conspecific (Engh et al. 2006). An individual deciding not to produce any response can have important behavioural implications and should not be overlooked. Importantly, it may be difficult in practice to distinguish between legitimate decisions not to respond and failures to perceive signals or differences between them. Reliance on prior results from psychophysical experiments may be necessary to assess the likelihood of legitimate no responses. In terms of *what* behavioural response to produce, there may be graded decisions, for example about the magnitude of a response (e.g. graded signals given in response to calls of escalating territorial challengers; Wagner 1989b). Alternatively, decisions may involve selection of a particular response from a range of finite choices appropriate to the context, such as when subordinate chacma baboons respond to the grunts of dominants by producing a fear bark. Decisions about *when* to produce a response are always made on a continuous temporal scale. Nevertheless, responses to signals must be made close enough in time after the reception of the initiating signal so that too much time does not elapse for conspecifics to determine it was a response

to the initial signal (Miller et al. 2009b). As social environments become more dynamic and complex, so too do the problems of monitoring environmental events (e.g. other signallers in a communication network) to decide whether, what and when to respond.

Acoustic signalling by male frogs and New World monkeys illustrates the diversity of ways receivers' decisions about whether, what and when to respond can depend on monitoring dynamic social environments. In frogs, for example, decisions about whether to vocally defend a calling site against intrusion by another signaller or to adopt silent satellite behaviour can depend on the receiver's subordinate level categorization of the signaller's size and fighting ability (Wagner 1989a). Decisions about what to signal (e.g. sexual advertisement versus aggression) can depend on the perceived proximity and individual identity of nearby neighbours in the chorus; these decisions exhibit high levels of dynamic plasticity that allow male frogs to tailor their signalling behaviour according to spatial and temporal variation in local social environments (e.g. Brenowitz & Rose 1994; Bee & Gerhardt 2001a; Humfeld et al. 2009). A key process determining when a male frog calls may involve an inhibitory resetting mechanism by which his intrinsic calling rhythm is altered or reset by the perception of a neighbour's calls (Greenfield 2005). The inhibitory control over the timing of signal production in frogs is not limited to hearing only the nearest neighbour. For example, in playback experiments with túngara frogs (*E. pustulosus*), Greenfield & Rand (2000) showed that decisions about when to vocalize depended on a set of rules governed by selective attention to a subset of nearby calling males.

Recent work on common marmosets, *Callithrix jacchus*, also illustrates how various processes of decision making occur during acoustic communication. Visually occluded marmosets produce species-specific long distance calls known as 'phee's'. Typically, conspecifics produce a phee in response to hearing one, a vocal behaviour known as an antiphonal call. As this interactive behaviour involves long sequences of alternating antiphonal calls, it reflects the fluid role reversal between signaller and receiver that can occur within individuals. Upon hearing an initiating phee call, marmosets must decide whether, what and when to produce the antiphonal response. Deciding whether to respond appears based on a combination of the social relationships of the individuals involved: (1) familiarity, as marmosets engage in longer antiphonal calling bouts with their cagemates than with other individuals (Miller & Wang 2006), and (2) the degree to which the other callers abide by the 'social rules'. Interactive playback experiments revealed that animals delaying their antiphonal call response beyond the socially permissible range are less likely to receive a subsequent antiphonal call from the other animal (Miller et al. 2009b). The latency of the antiphonal response seems to be an important cue for determining whether a particular conspecific is actively engaging in the vocal interaction with the individual. While all antiphonal calls in common marmosets are the species-typical phee call, decisions about the specific form this call will take (i.e. what to produce) are somewhat more complicated. Callers appear to plan specific spectral and temporal features of the vocalization prior to call production, such as the frequency contour and duration of each pulse (Miller et al. 2009a). The resulting acoustic structure may then have an impact on subsequent interactions. Decisions about the timing of when antiphonal calls are produced are significant because if the vocalization is to be considered an antiphonal call it must occur within a particular time window following the preceding call (Miller et al. 2009b). Within the socially acceptable time, more subtle changes in the latency to call may also carry communicative information. Miller & Wang (2006) found that the latency to call antiphonally depended on the social relationship of the two callers engaged in the interaction, suggesting that response latency is not arbitrary and may

communicate a specific message to the other caller. These studies of antiphonal calling in marmosets show how decision is not a unitary process, but can involve multiple components, each of which can influence the subsequent behaviour of receivers.

Summary

Decisions about whether, what and when to respond to a received signal must be made for any communication behaviour to be initiated, especially in species for which communication is an inherently interactive act in which individuals frequently switch roles between signaller and receiver. For signallers in many species, producing a signal is not purely reflexive, but involves dynamic decision processes after receiving signals in constantly changing social environments. The sophistication and control of these processes are sure to vary taxonomically. Given the dynamic, interactive nature of communication in many systems, the perceptual and cognitive processes involved in decision making by receivers represent key linkages between sensory transduction and motor output and, as such, are critical features of receivers' psychological landscapes that should be investigated more thoroughly in future comparative studies.

CONCLUSIONS

The emergence of receiver psychology some 20 years ago was prompted, in part, by a need to explain the bewildering diversity of signal forms found in nature, and in part as a response to previous emphases on the strategic (as opposed to tactical) designs of signals (Guilford & Dawkins 1991, 1993). The receiver psychology paradigm has led research on animal communication to uncover how diversity in signal design may be influenced by selection in response to psychological mechanisms in receivers that influence a signal's detectability, discriminability and memorability. What about the potential diversity of mechanisms for receiving those signals? An important priority for future research conducted under the banner of receiver psychology should be to investigate the evolution of receivers' psychological landscapes. This will entail identifying the proximate mechanisms, functional consequences and evolutionary histories associated with similarities and differences in perceptual and cognitive solutions to shared problems across different lineages. We concede this is no simple challenge. The elaborate and showy signals that draw so many scientists and casual observers to animal behaviour may turn out to be relatively easy to study compared to the internal perceptual and cognitive mechanisms by which receivers process them.

All of the perceptual and cognitive processes discussed above, and others not discussed here (Gentner & Margoliash 2002), represent potential features of a receiver's psychological landscape that play critical functional roles in communication. In keeping with the original scope of receiver psychology, many open questions remain concerning how the perceptual and cognitive processes discussed above might influence the evolution of signal design. Answering these questions should also be an important goal for future studies within the receiver psychology framework. However, if we are to understand the evolution of animal communication, we must also understand the mechanisms underlying the perceptual and cognitive processing of signals, as well as their evolution. Any paradigm placing emphasis predominantly on signal evolution and signaller behaviour will necessarily tell only part of the story of animal communication.

Broadening the notion of a receiver's psychological landscape to include the full range of perceptual and cognitive processes involved in receiving communication signals bridges traditional boundaries between studies of cognitive aspects of animal

behaviour (e.g. Griffin 1984; Allen & Bekoff 1997; Dukas & Ratcliffe 2009; Shettleworth 2009) and neuroethological studies of communication and sensory biology (e.g. Gentner & Margoliash 2002; Gerhardt & Huber 2002; Romanski & Ghazanfar 2009; Miller & Cohen 2010). Our goal in discussing experimental studies of a diversity of taxa was to illustrate that many of the basic problems experienced by receivers are widely shared. Indeed, much of our initial motivation for writing this essay arose because the two authors study communication in two very different and distantly related taxa (C.T.M., monkeys; M.A.B., frogs). We believe much can be learned about the proximate mechanisms underlying acoustic communication and their evolution by investigating potentially different solutions to common problems of signal reception in diverse taxonomic groups (e.g. Gerhardt & Huber 2002). Looking forward, we submit that rigorous comparative approaches within a broadened receiver psychology paradigm could, over the coming decades, reveal a great deal about the mechanisms and evolution of the perceptual and cognitive processes composing receivers' psychological landscapes.

Acknowledgments

We thank Dorothy Cheney, Tim Gentner, Carl Gerhardt, Eileen Hebets, Robert Seyfarth, Dave Stephens, members of the Miller and Bee labs, and two anonymous referees for critical comments on previous versions of the manuscript, and Teresa Nick for the recording of zebra finch song depicted in Figs 1 and 4. During the preparation of this manuscript, C.T.M. was supported by a grant from the National Institute on Deafness and Other Communication Disorders (NIDCD, R00-DC009007) and M.A.B. was supported by grants from the NIDCD (R03-DC008396 and R01-DC009582), the National Science Foundation (INT-0107304 and IOS-0842759), the International Graduate School for Neurosensory Science and Systems (Deutsche Forschungsgemeinschaft GK 591), the University of Minnesota Graduate School, and the McKnight Foundation.

References

- Akre, K. L. & Ryan, M. J. 2010. Complexity increases working memory for mating signals. *Current Biology*, **20**, 502–505.
- Alais, D. & Burr, D. 2004. The ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, **14**, 257–262.
- Alder, T. B. & Rose, G. J. 1998. Long-term temporal integration in the anuran auditory system. *Nature Neuroscience*, **1**, 519–523.
- Allen, C. & Bekoff, M. 1997. *Species of Mind*. Cambridge, Massachusetts: MIT Press.
- Aubin, T. & Jouventin, P. 2002. How to vocally identify kin in a crowd: the penguin model. *Advances in the Study of Behavior*, **31**, 243–277.
- Baddeley, A. 2003. Working memory: looking back and looking forward. *Nature Reviews Neuroscience*, **4**, 829–839.
- Balcombe, J. P. 1990. Vocal recognition of pups by mother Mexican free-tailed bats, *Tadarida brasiliensis mexicana*. *Animal Behaviour*, **39**, 960–966.
- Baugh, A. T., Akre, K. L. & Ryan, M. J. 2008. Categorical perception of a natural, multivariate signal: mating call recognition in tungara frogs. *Proceedings of the National Academy of Sciences, U.S.A.*, **105**, 8985–8988.
- Bee, M. A. 2006. Individual recognition in animal species. In: *Encyclopedia of Language and Linguistics* (Ed. by M. Naguib), pp. 617–626. New York: Elsevier.
- Bee, M. A. 2007. Sound source segregation in grey treefrogs: spatial release from masking by the sound of a chorus. *Animal Behaviour*, **74**, 549–558.
- Bee, M. A. 2008. Finding a mate at a cocktail party: spatial release from masking improves acoustic mate recognition in grey treefrogs. *Animal Behaviour*, **75**, 1781–1791.
- Bee, M. A. 2010. Spectral preferences and the role of spatial coherence in simultaneous integration in gray treefrogs (*Hyla chrysoscelis*). *Journal of Comparative Psychology*, **124**, 412–424.
- Bee, M. A. & Gerhardt, H. C. 2001a. Neighbour–stranger discrimination by territorial male bullfrogs (*Rana catesbeiana*): II. Perceptual basis. *Animal Behaviour*, **62**, 1141–1150.
- Bee, M. A. & Gerhardt, H. C. 2001b. Habituation as a mechanism of reduced aggression between neighboring territorial male bullfrogs (*Rana catesbeiana*). *Journal of Comparative Psychology*, **115**, 68–82.
- Bee, M. A. & Gerhardt, H. C. 2002. Individual voice recognition in a territorial frog (*Rana catesbeiana*). *Proceedings of the Royal Society B*, **269**, 1443–1448.

- Bee, M. A. & Klump, G. M. 2004. Primitive auditory stream segregation: a neurophysiological study in the songbird forebrain. *Journal of Neurophysiology*, **92**, 1088–1104.
- Bee, M. A. & Klump, G. M. 2005. Auditory stream segregation in the songbird forebrain: effects of time intervals on responses to interleaved tone sequences. *Brain, Behavior and Evolution*, **66**, 197–214.
- Bee, M. A. & Michéyl, C. 2008. The cocktail party problem: what is it? How can it be solved? And why should animal behaviorists study it? *Journal of Comparative Psychology*, **122**, 235–251.
- Bee, M. A. & Riemersma, K. K. 2008. Does common spatial origin promote the auditory grouping of temporally separated signal elements in grey treefrogs? *Animal Behaviour*, **76**, 831–843.
- Bee, M. A., Michéyl, C., Oxenham, A. J. & Klump, G. M. 2010. Neural adaptation to tone sequences in the songbird forebrain: patterns, determinants, and relation to the build-up of auditory streaming. *Journal of Comparative Physiology A*, **196**, 543–557.
- Bergman, T. J., Beehner, J. C., Cheney, D. L. & Seyfarth, R. M. 2003. Hierarchical classification by rank and kinship in baboons. *Science*, **302**, 1234–1236.
- Bird, J. & Darwin, C. J. 1998. Effects of a difference in fundamental frequency in separating two sentences. In: *Psychophysical and Physiological Advances in Hearing* (Ed. by A. R. Palmer, A. Rees, A. Q. Summerfield & R. Meddis), pp. 263–269. London: Whurr.
- Braaten, R. E. & Leary, J. C. 1999. Temporal induction of missing birdsong segments in European starlings. *Psychological Science*, **10**, 162–166.
- Bradbury, J. W. & Vehrencamp, S. L. 2011. *Principles of Animal Communication*. 2nd edn. Sunderland, Massachusetts: Sinauer.
- Branstetter, B. K. & Finneran, J. J. 2008. Comodulation masking release in bottlenose dolphins (*Tursiops truncatus*). *Journal of the Acoustical Society of America*, **124**, 625–633.
- Bregman, A. S. 1990. *Auditory Scene Analysis: the Perceptual Organization of Sound*. Cambridge, Massachusetts: MIT Press.
- Brenowitz, E. A. & Rose, G. J. 1994. Behavioral plasticity mediates aggression in choruses of the Pacific treefrog. *Animal Behaviour*, **47**, 633–641.
- Brokx, J. P. L. & Nootboom, S. G. 1982. Intonation and the perceptual separation of simultaneous voices. *Journal of Phonetics*, **10**, 23–36.
- Bronkhorst, A. W. 2000. The cocktail party phenomenon: a review of research on speech intelligibility in multiple-talker conditions. *Acustica*, **86**, 117–128.
- Brooks, R. J. & Falls, J. B. 1975. Individual recognition by song in white-throated sparrows. III. Song features used in individual recognition. *Canadian Journal of Zoology*, **53**, 1749–1761.
- Brumm, H. & Slabbekoorn, H. 2005. Acoustic communication in noise. *Advances in the Study of Behavior*, **35**, 151–209.
- Candolin, U. 2003. The use of multiple cues in mate choice. *Biological Reviews*, **78**, 575–595.
- Carlyon, R. P. 2004. How the brain separates sounds. *Trends in Cognitive Sciences*, **8**, 465–471.
- Catchpole, C. K. & Slater, P. J. B. 2008. *Bird Song: Biological Themes and Variations*. Cambridge: Cambridge University Press.
- Cheney, D. L., Seyfarth, R. M. & Silk, J. B. 1995. The responses of female baboons (*Papio cynocephalus ursinus*) to anomalous social interactions: evidence for causal reasoning. *Journal of Comparative Psychology*, **109**, 134–141.
- Cherry, E. C. 1953. Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America*, **25**, 975–979.
- Christensen-Dalsgaard, J. & Carr, C. E. 2008. Evolution of a sensory novelty: tympanic ears and the associated neural processing. *Brain Research Bulletin*, **75**, 365–370.
- Clucas, B. A., Freeberg, T. M. & Lucas, J. R. 2004. Chick-a-dee call syntax, social context, and season affect vocal responses of Carolina chickadees (*Parus carolinensis*). *Behavioral Ecology and Sociobiology*, **57**, 187–196.
- Comins, J. A. & Gentner, T. Q. 2010. Working memory for patterned sequences of auditory objects in a songbird. *Cognition*, **117**, 38–53.
- Dapper, A. L., Baugh, A. T. & Ryan, M. J. 2011. The sounds of silence as an alarm cue in túngara frogs, *Physalaemus pustulosus*. *Biotropica*, **43**, 380–385.
- Darwin, C. J. 1997. Auditory grouping. *Trends in Cognitive Sciences*, **1**, 327–333.
- Darwin, C. J. 2008. Listening to speech in the presence of other sounds. *Philosophical Transactions of the Royal Society B*, **363**, 1011–1021.
- Darwin, C. J. & Carlyon, R. P. 1995. Auditory grouping. In: *Hearing* (Ed. by B. C. J. Moore), pp. 387–424. New York: Academic Press.
- Dent, M. L., McClaine, E. M., Best, V., Ozmeral, E., Narayan, R., Gallun, F. J., Sen, K. & Shinn-Cunningham, B. G. 2009. Spatial unmasking of birdsong in zebra finches (*Taeniopygia guttata*) and budgerigars (*Melopsittacus undulatus*). *Journal of Comparative Psychology*, **123**, 357–367.
- Dukas, R. & Ratcliffe, J. M. 2009. *Cognitive Ecology II*. Chicago: University of Chicago Press.
- Edwards, C. J., Alder, T. B. & Rose, G. J. 2002. Auditory midbrain neurons that count. *Nature Neuroscience*, **5**, 934–936.
- Engh, A. L., Hoffmeier, R. R., Cheney, D. L. & Seyfarth, R. M. 2006. Who, me? Can baboons infer the target of vocalizations? *Animal Behaviour*, **71**, 381–387.
- Farris, H. E. & Ryan, M. J. 2011. Relative comparisons of call parameters enable auditory grouping in frogs. *Nature Communications*, **2**, 410.
- Farris, H. E., Rand, A. S. & Ryan, M. J. 2002. The effects of spatially separated call components on phonotaxis in túngara frogs: evidence for auditory grouping. *Brain Behavior and Evolution*, **60**, 181–188.
- Farris, H. E., Rand, A. S. & Ryan, M. J. 2005. The effects of time, space and spectrum on auditory grouping in túngara frogs. *Journal of Comparative Physiology A*, **191**, 1173–1183.
- Fay, R. R. 1998. Auditory stream segregation in goldfish (*Carassius auratus*). *Hearing Research*, **120**, 69–76.
- Fay, R. R. 2011. Signal-to-noise ratio for source determination and for a comodulated masker in goldfish, *Carassius auratus*. *Journal of the Acoustical Society of America*, **129**, 3367–3372.
- Feng, A. S. & Ratnam, R. 2000. Neural basis of hearing in real-world situations. *Annual Review of Psychology*, **51**, 699–725.
- Fishman, Y. I., Reser, D. H., Arezzo, J. C. & Steinschneider, M. 2001. Neural correlates of auditory stream segregation in primary auditory cortex of the awake monkey. *Hearing Research*, **151**, 167–187.
- Fishman, Y. I., Arezzo, J. C. & Steinschneider, M. 2004. Auditory stream segregation in monkey auditory cortex: effects of frequency separation, presentation rate, and tone duration. *Journal of the Acoustical Society of America*, **116**, 1656–1670.
- Freeberg, T. M. & Lucas, J. R. 2002. Receivers respond differently to chick-a-dee calls varying in note composition in Carolina chickadees, *Parus carolinensis*. *Animal Behaviour*, **63**, 837–845.
- Füllgrabe, C., Berthommier, F. & Lorenzi, C. 2006. Masking release for consonant features in temporally fluctuating background noise. *Hearing Research*, **211**, 74–84.
- Geissler, D. B. & Ehret, G. 2002. Time-critical integration of formants for perception of communication calls in mice. *Proceedings of the National Academy of Sciences, U.S.A.*, **99**, 9021–9025.
- Gentner, T. Q. 2004. Neural systems for individual song recognition in adult birds. In: *Behavioral Neurobiology of Birds* (Ed. by H. P. Zeigler & P. Marler), pp. 282–302. New York: New York Academy of Sciences.
- Gentner, T. Q. 2008. Temporal scales of auditory objects underlying birdsong vocal recognition. *Journal of the Acoustical Society of America*, **124**, 1350–1359.
- Gentner, T. Q. & Hulse, S. H. 1998. Perceptual mechanisms for individual vocal recognition in European starlings, *Sturnus vulgaris*. *Animal Behaviour*, **56**, 579–594.
- Gentner, T. Q. & Hulse, S. H. 2000. Perceptual classification based on the component structure of song in European starlings. *Journal of the Acoustical Society of America*, **107**, 3369–3381.
- Gentner, T. Q. & Margoliash, D. 2002. The neuroethology of vocal communication: perception and cognition. In: *Acoustic Communication* (Ed. by A. M. Simmons, R. R. Fay & A. N. Popper), pp. 324–386. New York: Springer.
- Gerhardt, H. C. 1992. Multiple messages in acoustic signals. *Seminars in the Neurosciences*, **4**, 391–400.
- Gerhardt, H. C. 2001. Acoustic communication in two groups of closely related treefrogs. *Advances in the Study of Behavior*, **30**, 99–167.
- Gerhardt, H. C. & Huber, F. 2002. *Acoustic Communication in Insects and Anurans: Common Problems and Diverse Solutions*. Chicago: Chicago University Press.
- Gerhardt, H. C., Roberts, J. D., Bee, M. A. & Schwartz, J. J. 2000. Call matching in the quacking frog (*Crinia georgiana*). *Behavioral Ecology and Sociobiology*, **48**, 243–251.
- Ghazanfar, A. A. & Logothetis, N. K. 2003. Neuroperception: facial expressions linked to monkey calls. *Nature*, **423**, 937–938.
- Greenfield, M. D. 2005. Mechanisms and evolution of communal sexual displays in arthropods and anurans. *Advances in the Study of Behavior*, **35**, 1–62.
- Greenfield, M. D. & Rand, A. S. 2000. Frogs have rules: selective attention algorithms regulate chorusing in *Physalaemus pustulosus* (Leptodactylidae). *Ethology*, **106**, 331–347.
- Griffin, D. R. 1984. *Animal Thinking*. Cambridge, Massachusetts: Harvard University Press.
- Griffiths, T. D. & Warren, J. D. 2004. What is an auditory object? *Nature Reviews Neuroscience*, **5**, 887–892.
- Guilford, T. & Dawkins, M. S. 1991. Receiver psychology and the evolution of animal signals. *Animal Behaviour*, **42**, 1–14.
- Guilford, T. & Dawkins, M. S. 1993. Receiver psychology and the design of animal signals. *Trends in Neurosciences*, **16**, 430–436.
- Hartmann, W. M. 1988. Pitch perception and the segregation and integration of auditory entities. In: *Auditory Function: Neurobiological Basis of Hearing* (Ed. by G. M. Edelman, W. E. Gall & W. M. Cowan), pp. 623–645. New York: J. Wiley.
- Heil, P. & Neubauer, H. 2003. A unifying basis of auditory thresholds based on temporal summation. *Proceedings of the National Academy of Sciences, U.S.A.*, **100**, 6151–6156.
- Hulse, S. H. 2002. Auditory scene analysis in animal communication. *Advances in the Study of Behavior*, **31**, 163–200.
- Humfeld, S. C., Marshall, V. T. & Bee, M. A. 2009. Context-dependent plasticity of aggressive signalling in a dynamic social environment. *Animal Behaviour*, **78**, 915–924.
- Izumi, A. 2002. Auditory stream segregation in Japanese monkeys. *Cognition*, **82**, B113–B122.
- Izumi, A. & Kojima, S. 2004. Matching vocalizations to vocalizing faces in a chimpanzee (*Pan troglodytes*). *Animal Cognition*, **7**, 179–184.
- Jensen, K. K. 2007. Comodulation detection differences in the hooded crow (*Corvus corone corone*), with direct comparison to human subjects. *Journal of the Acoustical Society of America*, **121**, 1783–1789.
- King, A. J. 2007. Auditory neuroscience: filling in the gaps. *Current Biology*, **17**, R799–R801.
- Klump, G. M. 1996. Bird communication in the noisy world. In: *Ecology and Evolution of Acoustic Communication in Birds* (Ed. by D. E. Kroodsma & E. H. Miller), pp. 321–338. Ithaca, New York: Cornell University Press.
- Klump, G. M. & Langemann, U. 1995. Comodulation masking release in a songbird. *Hearing Research*, **87**, 157–164.

- Knudsen, D., Thompson, J. V. & Gentner, T. Q. 2010. Distributed recognition of natural songs in European starlings. *Learning & Motivation*, **41**, 287–306.
- Kwon, B. J. 2002. Comodulation masking release in consonant recognition. *Journal of the Acoustical Society of America*, **112**, 634–641.
- Langemann, U. & Klump, G. M. 2001. Signal detection in amplitude-modulated maskers. I. Behavioural auditory thresholds in a songbird. *European Journal of Neuroscience*, **13**, 1025–1032.
- Langemann, U. & Klump, G. M. 2005. Perception and acoustic communication networks. In: *Animal Communication Networks* (Ed. by P. K. McGregor), pp. 451–480. Cambridge: Cambridge University Press.
- Langemann, U. & Klump, G. M. 2007. Detecting modulated signals in modulated noise: I. Behavioural auditory thresholds in a songbird. *European Journal of Neuroscience*, **26**, 1969–1978.
- Lee, D. 2010. Neuroethology of decision making. In: *Primate Neuroethology* (Ed. by M. Platt & A. A. Ghazanfar), pp. 550–569. New York: Oxford University Press.
- Ma, L., Micheyl, C., Yin, P. B., Oxenham, A. J. & Shamma, S. A. 2010. Behavioral measures of auditory streaming in ferrets (*Mustela putorius*). *Journal of Comparative Psychology*, **124**, 317–330.
- MacDougall-Shackleton, S. A., Hulse, S. H., Gentner, T. Q. & White, W. 1998. Auditory scene analysis by European starlings (*Sturnus vulgaris*): perceptual segregation of tone sequences. *Journal of the Acoustical Society of America*, **103**, 3581–3587.
- McDermott, J. H. 2009. The cocktail party problem. *Current Biology*, **19**, R1024–R1027.
- McGregor, P. K. 2005. *Animal Communication Networks*. Cambridge: Cambridge University Press.
- McGurk, H. & Macdonald, J. 1976. Hearing lips and seeing voices. *Nature*, **264**, 746–748.
- Miller, C. T. & Cohen, Y. E. 2010. Vocalizations as auditory objects: behavior and neurophysiology. In: *Primate Neuroethology* (Ed. by M. Platt & A. Ghazanfar), pp. 236–254. : Oxford University Press.
- Miller, C. T. & Wang, X. 2006. Sensory-motor interactions modulate a primate vocal behavior: antiphonal calling in common marmosets. *Journal of Comparative Physiology A*, **192**, 27–38.
- Miller, C. T., Dibble, E. & Hauser, M. D. 2001. Amodal completion of acoustic signals by a nonhuman primate. *Nature Neuroscience*, **4**, 783–784.
- Miller, C. T., Scarl, J. S. & Hauser, M. D. 2004. Sex-specific sensory biases underlie sex differences in tamarin long call structure. *Animal Behaviour*, **68**, 713–720.
- Miller, C. T., Iguina, C. G. & Hauser, M. D. 2005. Processing vocal signals for recognition during antiphonal calling in tamarins. *Animal Behaviour*, **69**, 1387–1398.
- Miller, C. T., Eliades, S. J. & Wang, X. 2009a. Motor-planning for vocal production in common marmosets. *Animal Behaviour*, **78**, 1195–1203.
- Miller, C. T., Beck, K., Meade, B. & Wang, X. 2009b. Antiphonal call timing in marmosets is behaviorally significant: interactive playback experiments. *Journal of Comparative Physiology A*, **195**, 783–789.
- Mitani, J. & Marler, P. 1989. A phonological analysis of male gibbon singing behavior. *Behaviour*, **109**, 20–45.
- Møller, A. R. 2003. *Sensory Systems: Anatomy and Physiology*. Amsterdam: Academic Press.
- Myrberg, A. A. & Riggio, R. J. 1985. Acoustically mediated individual recognition by a coral reef fish (*Pomacentrus partitus*). *Animal Behaviour*, **33**, 411–416.
- Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P. & Winkler, I. 2001. 'Primitive intelligence' in the auditory cortex. *Trends in Neurosciences*, **24**, 283–288.
- Narins, P. M., Grabul, D. S., Soma, K. K., Gaucher, P. & Hödl, W. 2005. Cross-modal integration in a dart-poison frog. *Proceedings of the National Academy of Sciences, U.S.A.*, **102**, 2425–2429.
- Nelken, I., Rotman, Y. & Bar Yosef, O. 1999. Responses of auditory-cortex neurons to structural features of natural sounds. *Nature*, **397**, 154–157.
- Nelson, D. A. 1988. Feature weighting in species song recognition by the field sparrow (*Spizella pusilla*). *Behaviour*, **106**, 158–182.
- Nelson, D. A. 1989. Song frequency as a cue for recognition of species and individuals in the field sparrow (*Spizella pusilla*). *Journal of Comparative Psychology*, **103**, 171–176.
- Nityananda, V. & Bee, M. A. 2011. Finding your mate at a cocktail party: frequency separation promotes auditory stream segregation of concurrent voices in multi-species frog choruses. *PLoS ONE*, **6**, e21191.
- Nityananda, V. & Bee, M. A. In press. Spatial release from masking in a free-field source identification task by gray treefrogs. *Hearing Research*.
- van Noorden, L. P. A. S. 1975. Temporal Coherence in the Perception of Tone Sequences. Ph.D. thesis, Eindhoven University of Technology.
- Payne, K. B. & Payne, R. S. 1985. Large scale changes over 19 years in songs of humpback whales in Bermuda. *Zeitschrift für Tierpsychologie*, **68**, 89–114.
- Peake, T. M. 2005. Eavesdropping in communication networks. In: *Animal Communication Networks* (Ed. by P. K. McGregor), pp. 13–37. Cambridge: Cambridge University Press.
- Petkov, C. I., O'Connor, K. N. & Sutter, M. L. 2003. Illusory sound perception in macaque monkeys. *Journal of Neuroscience*, **23**, 9155–9161.
- Pressnitzer, D., Sayles, M., Micheyl, C. & Winter, I. M. 2008. Perceptual organization of sound begins in the auditory periphery. *Current Biology*, **18**, 1124–1128.
- Proops, L., McComb, K. & Reby, D. 2009. Cross-modal individual recognition in domestic horses (*Equus caballus*). *Proceedings of the National Academy of Sciences, U.S.A.*, **106**, 947–951.
- Ranganath, C. & Rainer, G. 2003. Neural mechanisms for detecting and remembering novel events. *Nature Reviews Neuroscience*, **4**, 193–202.
- Rendall, D., Rodman, P. S. & Emond, R. E. 1996. Vocal recognition of individuals and kin in free-ranging rhesus monkeys. *Animal Behaviour*, **51**, 1007–1015.
- Richardson, C. & Lengagne, T. 2010. Multiple signals and male spacing affect female preference at cocktail parties in treefrogs. *Proceedings of the Royal Society B*, **277**, 1247–1252.
- Robinson, J. G. 1979. An analysis of the organization of vocal communication in the titi monkey *Callicebus moloch*. *Zeitschrift für Tierpsychologie*, **49**, 381–405.
- Romanski, L. M. & Ghazanfar, A. A. 2009. The primate frontal and temporal lobes and their role in multisensory vocal communication. In: *Primate Neuroethology* (Ed. by M. L. Platt & A. A. Ghazanfar), pp. 500–524. Oxford: Oxford University Press.
- Ronacher, B. & Hoffmann, C. 2003. Influence of amplitude modulated noise on the recognition of communication signals in the grasshopper *Chorthippus biguttulus*. *Journal of Comparative Physiology A*, **189**, 419–425.
- Rosch, E. 1978. Principles of categorization. In: *Cognition and Categorization* (Ed. by E. Rosch & B. B. Lloyd), pp. 27–48. Hillsdale, New Jersey: L. Erlbaum.
- Rosch, E., Mervis, C. B., Gray, W. B., Johnson, D. M. & Boyes-Braem, P. 1976. Basic objects in natural categories. *Cognitive Psychology*, **8**, 349–382.
- Rose, G. J. & Brenowitz, E. A. 2002. Pacific treefrogs use temporal integration to differentiate advertisement from encounter calls. *Animal Behaviour*, **63**, 1183–1190.
- Rose, M. M. & Moore, B. C. J. 2005. The relationship between stream segregation and frequency discrimination in normally hearing and hearing-impaired subjects. *Hearing Research*, **204**, 16–28.
- Rosenthal, G. G., Rand, A. S. & Ryan, M. J. 2004. The vocal sac as a visual cue in anuran communication: an experimental analysis using video playback. *Animal Behaviour*, **68**, 55–58.
- Rowe, C. 1999. Receiver psychology and the evolution of multicomponent signals. *Animal Behaviour*, **58**, 921–931.
- Rowe, C. & Guilford, T. 1999. The evolution of multimodal warning displays. *Evolutionary Ecology*, **13**, 655–671.
- Rowe, C. & Skellhorn, J. 2004. Avian psychology and communication. *Proceedings of the Royal Society B*, **271**, 1435–1442.
- Samuel, A. G. 1996. Does lexical information influence the perceptual restoration of phonemes? *Journal of Experimental Psychology: General*, **125**, 28–51.
- Sayigh, L. S., Tyack, P. L., Wells, R. S., Solow, A. R., Scott, M. D. & Irvine, A. B. 1999. Individual recognition in wild bottlenose dolphins: a field test using playback experiments. *Animal Behaviour*, **57**, 41–50.
- Schleidt, W. M. 1973. Tonic communication: continual effects of discrete signs in animal communication systems. *Journal of Theoretical Biology*, **42**, 359–386.
- Scholl, B. J. 2001. Objects and attention: the state of the art. *Cognition*, **80**, 1–46.
- Scholl, B. J. & Leslie, A. M. 1999. Explaining the infant's object concept: beyond the perception/cognition dichotomy. In: *What is Cognitive Science?* (Ed. by E. Lepore & Z. Pylyshyn), pp. 26–73 Oxford: Blackwell Press.
- Schul, J. & Sheridan, R. A. 2006. Auditory stream segregation in an insect. *Neuroscience*, **138**, 1–4.
- Schwartz, J. J. & Gerhardt, H. C. 1989. Spatially mediated release from auditory masking in an anuran amphibian. *Journal of Comparative Physiology A*, **166**, 37–41.
- Schwartz, J. J. & Gerhardt, H. C. 1995. Directionality of the auditory system and call pattern recognition during acoustic interference in the gray treefrog, *Hyla versicolor*. *Auditory Neuroscience*, **1**, 195–206.
- Schwartz, J. J., Huth, K., Jones, S. H., Brown, R. & Marks, J. 2010. Tests for call restoration in the gray treefrog, *Hyla versicolor*. *Bioacoustics*, **20**, 59–86.
- Searcy, W. A. 1983. Response to multiple song types in the male song sparrows and field sparrows. *Animal Behaviour*, **31**, 948–949.
- Searcy, W. A. & Nowicki, S. 1999. Functions of song variation in song sparrows. In: *The Design of Animal Communication* (Ed. by M. D. Hauser & M. Konishi), pp. 577–596. Cambridge, Massachusetts: MIT Press.
- Searcy, W. A., Podos, J., Peters, S. & Nowicki, S. 1995. Discrimination of song types and variants in song sparrows. *Animal Behaviour*, **49**, 1219–1226.
- Searcy, W. A., Nowicki, S. & Peters, S. 1999. Song types as fundamental units in vocal repertoires. *Animal Behaviour*, **58**, 37–44.
- Seeba, F. & Klump, G. M. 2009. Stimulus familiarity affects perceptual restoration in the European starling (*Sturnus vulgaris*). *PLoS ONE*, **4**, e5974.
- Seeba, F., Schwartz, J. J. & Bee, M. A. 2010. Testing an auditory illusion in frogs: perceptual restoration or sensory bias? *Animal Behaviour*, **79**, 1317–1328.
- Shettleworth, S. J. 2009. *Cognition, Evolution, and Behavior*. Oxford: Oxford University Press.
- Shinn-Cunningham, B. G., Ihlefeld, A., Satyavarta & Larson, E. 2005. Bottom-up and top-down influences on spatial unmasking. *Acta Acustica United with Acustica*, **91**, 967–979.
- Simmons, A. M. & Bean, M. E. 2000. Perception of mistuned harmonics in complex sounds by the bullfrog (*Rana catesbeiana*). *Journal of Comparative Psychology*, **114**, 167–173.
- Simmons, A. M., Popper, A. N. & Fay, R. R. 2003. *Acoustic Communication*. New York: Springer-Verlag.
- Simons, D. J. 1996. In sight, out of mind: when object representations fail. *Psychological Science*, **7**, 301–305.
- Sliwa, J., Duhamel, J. R., Pascalis, O. & Wirth, S. 2011. Spontaneous voice-face identity matching by rhesus monkeys for familiar conspecifics and humans. *Proceedings of the National Academy of Sciences, U.S.A.*, **108**, 1735–1740.
- Sloan, J. L. & Hare, J. F. 2004. Monotony and the information content of Richardson's ground squirrel (*Spermophilus richardsonii*) repeated calls: tonic communication or signal certainty? *Ethology*, **110**, 147–156.
- Snyder, J. S. & Alain, C. 2007. Toward a neurophysiological theory of auditory stream segregation. *Psychological Bulletin*, **133**, 780–799.

- Speed, M. P. 2000. Warning signals, receiver psychology and predator memory. *Animal Behaviour*, **60**, 269–278.
- Spelke, E. S., Breinlinger, K., Jacobson, K. & Phillips, A. 1993. Gestalt relations and object perception: a developmental study. *Perception*, **22**, 1483–1501.
- Stokes, A. 1967. Behavior of the bobwhite, *Colinus virginianus*. *Auk*, **84**, 1–33.
- Sumby, W. H. & Pollack, I. 1954. Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America*, **26**, 212–215.
- Taylor, R. C., Klein, B. A., Stein, J. & Ryan, M. J. 2008. Faux frogs: multimodal signalling and the value of robotics in animal behaviour. *Animal Behaviour*, **76**, 1089–1097.
- Taylor, R. C., Klein, B. A., Stein, J. & Ryan, M. J. 2011. Multimodal signal variation in space and time: how important is matching a signal with its signaler? *Journal of Experimental Biology*, **214**, 815–820.
- ten Cate, C. & Rowe, C. 2007. Biases in signal evolution: learning makes a difference. *Trends in Ecology & Evolution*, **22**, 380–387.
- Vélez, A. & Bee, M. A. 2010. Signal recognition by frogs in the presence of temporally fluctuating chorus-shaped noise. *Behavioral Ecology and Sociobiology*, **64**, 1695–1709.
- Vélez, A. & Bee, M. A. 2011. Dip listening and the cocktail party problem in grey treefrogs: signal recognition in temporally fluctuating noise. *Animal Behaviour*, **82**, 1319–1327.
- Verhey, J. L., Pressnitzer, D. & Winter, I. M. 2003. The psychophysics and physiology of comodulation masking release. *Experimental Brain Research*, **153**, 405–417.
- Wagner, W. E. 1989a. Fighting, assessment, and frequency alteration in Blanchard cricket frog. *Behavioral Ecology and Sociobiology*, **25**, 429–436.
- Wagner, W. E. 1989b. Graded aggressive signals in Blanchard's cricket frog: vocal responses to opponent proximity and size. *Animal Behaviour*, **38**, 1025–1038.
- Warren, R. M. 1970. Perceptual restoration of missing speech sounds. *Science*, **167**, 392–393.
- Webster, D. B., Fay, R. R. & Popper, A. N. 1992. *The Evolutionary Biology of Hearing*. New York: Springer.
- Weiss, D. J. & Hauser, M. D. 2002. Perception of harmonics in the combination long call of cottontop tamarins, *Saguinus oedipus*. *Animal Behaviour*, **64**, 415–426.
- Yost, W. A., Popper, A. N. & Fay, R. R. 2008. *Auditory Perception of Sound Sources*. New York: Springer.
- Zokoll, M. A., Klump, G. M. & Langemann, U. 2007. Auditory short-term memory persistence for tonal signals in a songbird. *Journal of the Acoustical Society of America*, **121**, 2842–2851.
- Zokoll, M. A., Klump, G. M. & Langemann, U. 2008. Auditory memory for temporal characteristics of sound. *Journal of Comparative Physiology A*, **194**, 457–467.
- Zuberbühler, K. 2002. A syntactic rule in forest monkey communication. *Animal Behaviour*, **63**, 293–299.