The Cocktail Party Problem: What Is It? How Can It Be Solved? and Why Should Animal Behaviorists Study It?

Mark A. Bee and Christophe Micheyl
University of Minnesota–Twin Cities

Animals often use acoustic signals to communicate in groups or social aggregations in which multiple individuals signal within a receiver's hearing range. Consequently, receivers face challenges related to acoustic interference and auditory masking that are not unlike the human *cocktail party problem*, which refers to the problem of perceiving speech in noisy social settings. Understanding the sensory solutions to the cocktail party problem has been a goal of research on human hearing and speech communication for several decades. Despite a general interest in acoustic signaling in groups, animal behaviorists have devoted comparatively less attention toward understanding how animals solve problems equivalent to the human cocktail party problem. After illustrating how humans and nonhuman animals experience and overcome similar perceptual challenges in cocktail-party-like social environments, this article reviews previous psychophysical and physiological studies of humans and nonhuman animals to describe how the cocktail party problem can be solved. This review also outlines several basic and applied benefits that could result from studies of the cocktail party problem in the context of animal acoustic communication.

AQ: 1

Keywords:

In many animals, acoustic communication occurs in large groups or aggregations of signaling individuals (Gerhardt & Huber, 2002; Greenfield, 2005; Kroodsma & Miller, 1996; McGregor, 2005; this issue). It is important to appreciate that the sounds produced by multiple signalers in groups, as well as other biotic and abiotic sources of noise and acoustic reflections from objects in the environment, are sound pressure waves that add together to form a composite pressure waveform that impinges on a receivers' hearing organs. Thus, signaling in groups poses a number of special problems for receivers in terms of detecting and classifying signals (Brumm & Slabbekoorn, 2005; Hulse, 2002; Klump, 1996; Langemann & Klump, 2005; Wiley, 2006). These problems, and their solutions, are our concern here. For humans, solutions to these problems rest, in part, on the auditory system's ability to parse the composite acoustic waveform generated by multiple sources into perceptually coherent representations—termed *auditory objects*, *auditory images*, or *auditory streams*—that represent different sound sources in the acoustic scene (Bregman, 1990; Van Valkenburg & Kubovy, 2004; Yost, 1991).

One well-known example of a problem in human auditory scene analysis (Bregman, 1990) is the aptly named *cocktail party prob-lem* (Cherry, 1953), which refers to the difficulty we sometimes have understanding speech in noisy social settings (reviewed in Bronkhorst, 2000; Yost, 1997). In this article, we discuss issues relating to auditory scene analysis and the cocktail party problem that have a long history in studies of human hearing and speech communication but have received less attention in the study of animal acoustic communication. The take-home messages from this article are (a) that many of the sensory solutions to the human cocktail party problem represent potentially important mechanisms underlying acoustic communication in nonhuman animals and (b) that the study of these mechanisms has important implications for our understanding of animal acoustic communication.

This article is organized around three questions. In the first section, we pose the question, What is the cocktail party problem? Here, we more explicitly outline the issues at hand and show that nonhuman animals also encounter and solve cocktail-party-like problems. In the second section, we ask, How can the cocktail party problem be solved? In this section, we draw extensively on studies of human hearing and speech perception to illustrate the diversity of mechanisms that allow humans to perceive speech in noisy social settings. We also show how many of these same auditory mechanisms operate in nonhuman animals. These studies could serve as useful heuristics to guide future research on the mechanisms of animal acoustic communication. In the third and final section we ask, Why should animal behaviorists study the cocktail party problem? Here, we raise a number of basic and applied issues to suggest the potential benefits and opportunities associated with studies of animal "cocktail parties."

The fields of study we attempt to integrate are broad; therefore, the scope of our review is necessarily broad and we make no attempt to cover any single topic in great depth. For other treatments on these topics, readers are referred to Bregman (1990, 1993, Bronkhorst (2000), Carlyon (2004), Darwin (1997), and

AQ: 2

2 BEE AND MICHEYL

Darwin and Carlyon (1995) for work in humans and to Brumm and Slabbekoorn (2005), Feng and Ratnam (2000), Hulse (2002), Klump (1996, 2005), and Langemann and Klump (2005) for work in animals.

AQ: 3

What Is the Cocktail Party Problem?

The essence of the cocktail party problem can be formulated as a deceptively simple question: "How do we recognize what one person is saying when others are speaking at the same time?" (Bronkhorst, 2000, p. 117). Finding answers to this question has been an important goal of human hearing research for several decades (see Bronkhorst, 2000, and Yost, 1997, for reviews). At the root of the cocktail party problem is the fact that the human voices present in a noisy social setting often overlap in frequency and in time and thus represent sources of direct acoustic interference and *energetic masking* that can impair the perception of speech. In addition, recent research has revealed that even those components of concurrent speech that do not overlap in frequency or time with those of the target signal can dramatically affect speech intelligibility via so-called *informational masking* (Shinn-Cunningham, Ihlefeld, Satyavarta, & Larson, 2005). The ability of concurrent speech and speechlike noise to impair speech perception is well-documented in the literature on human hearing (reviewed in Bronkhorst, 2000).

Do Animals Experience Cocktail-Party-Like Problems?

Examples of nonhuman animals that acoustically communicate in groups or large social aggregations abound in the animal behavior literature. Among the best-known examples are perhaps frog and insect choruses, the songbird dawn chorus, and flocking and colonial birds (reviewed in Hulse, 2002). For animals that acoustically communicate in such groups, the problem of perceiving acoustic signals is equivalent to the human cocktail party problem because the signals of multiple conspecifics often occur concurrently (Brumm & Slabbekoorn, 2005; Hulse, 2002; Klump, 1996). As in humans, the consequences of interference and masking for other animals include increases in signal detection thresholds and decreases in the ability to recognize and discriminate among different signal variants (Bee, in press; Bee & Swanson, in press; Ehret & Gerhardt, 1980; Gerhardt & Klump, 1988; Langemann, Gauger, & Klump, 1998; Lohr, Wright, & Dooling, 2003; Schwartz & Gerhardt, 1989, 1995; Wollerman, 1999; Wollerman & Wiley, 2002). Hence, humans and nonhuman animals are perhaps not so different when it comes to the problems faced when acoustically communicating in groups. For animals, we can formalize the cocktail party problem in the following question: How do animals detect and recognize conspecific signals, localize signalers, discriminate among signal types and individual signalers, and extract information from signals and signaling interactions when multiple conspecifics and heterospecifics are signaling at the same time? Note that this formalization includes the contribution of heterospecific signalers to a receiver's cocktail party problem. For many animals, the signals of heterospecifics might compound the cocktail party problem in ways that have not been considered previously in the literature on human hearing and speech communication.

Despite evidence that humans and other animals encounter problems perceiving acoustic signals in noisy groups, the cocktail party problem is not insurmountable. Personal experience tells us this is the case for humans. Field observations and playback experiments indicate that this is also the case for nonhuman animals. Consider the following examples. During their breeding seasons, frogs commonly aggregate in mixed-species choruses in which males produce loud advertisement calls (Gerhardt, 1975), and the background noise levels in a chorus can be quite high (Narins, 1982; Narins & Zelick, 1988). Nevertheless, within a chorus environment, male advertisement calls can be used by females to make adaptive mate choice decisions and by males to mediate male–male aggressive interactions (reviewed in Gerhardt & Bee, 2006; Gerhardt & Huber, 2002; Wells & Schwartz, 2006). Like frogs, songbirds communicate in noisy situations, such as the dawn chorus, and use acoustic signals for mate attraction and territory defense (Klump, 1996). The work of Hulse and others has shown that songbirds can correctly recognize the songs of a particular species and of particular individuals when these songs are digitally mixed with the songs of other species or individuals and even when they are mixed with the sounds of a dawn chorus (Benney & Braaten, 2000; Hulse, MacDougall-Shackleton, & Wisniewski, 1997; Wisniewski & Hulse, 1997). Bank swallows, cliff swallows, and king penguins are three colonial bird species for which acoustically mediated parent–offspring recognition allows parents to reunite with chicks in a large and noisy colony (Aubin & Jouventin, 1998, 2002; Beecher, 1989, 1991). Field playback tests with king penguins have revealed that chicks can detect parental calls (the signal) even when these were mixed with the calls of five other adults (the noise) at a signal-to-noise ratio of -6 dB (Aubin & Jouventin, 1998). Clearly, a diversity of animals signal in groups, and these animals both encounter and solve cocktail-party-like problems.

How Can the Cocktail Party Problem Be Solved?

The main point we wish to emphasize in this article is that effective acoustic communication will often depend heavily on the perceptual mechanisms that receivers possess for solving cocktail-party-like problems. To be sure, for many animals, both the structure of acoustic signals and the behavior of signalers represent adaptations that have evolved as a result of selection pressures associated with ameliorating cocktail-party-like problems for receivers (reviewed in Brumm & Slabbekoorn, 2005; Klump, 2005; Wiley, 2006). In this section, we focus exclusively on the receiver side of things by reviewing literature related to the perceptual processes that contribute to solving the cocktail party problem. The cocktail party problem represents a specific example of the more general task of auditory scene analysis (Bregman, 1990), which refers to the processes that form coherent and functional perceptual representations of distinct sound sources in the environment (Bregman, 1990; Hulse, 2002). In this section, we review some important topics in auditory scene analysis and describe how they relate to animal acoustic communication. We then show how several processes that function in auditory scene analysis contribute to solving the cocktail party problem in humans, and we illustrate how these same mechanisms operate in nonhuman animals.

Auditory Scene Analysis

Sequential and simultaneous integration. In his book on auditory scene analysis, Bregman (1990) proposed an important distinction between sequential integration and simultaneous integration. *Sequential integration* refers to the integration of temporally separated sounds from one sound source (e.g., syllables, words; Figure 1A) into a coherent auditory stream and the segregation of these sounds from other intervening and overlapping sounds from other sources. *Simultaneous integration* refers to the perceptual grouping of different, simultaneously occurring components of the frequency spectrum (e.g., harmonics, speech formants; Figure 1A) into a representation of a single sound source and the segregation of these sounds from other concurrent sounds in the environment. Integration and segregation are often regarded as converse but complementary processes: When particular sound elements are integrated together, they are also segregated from other sounds (Bregman, 1990).
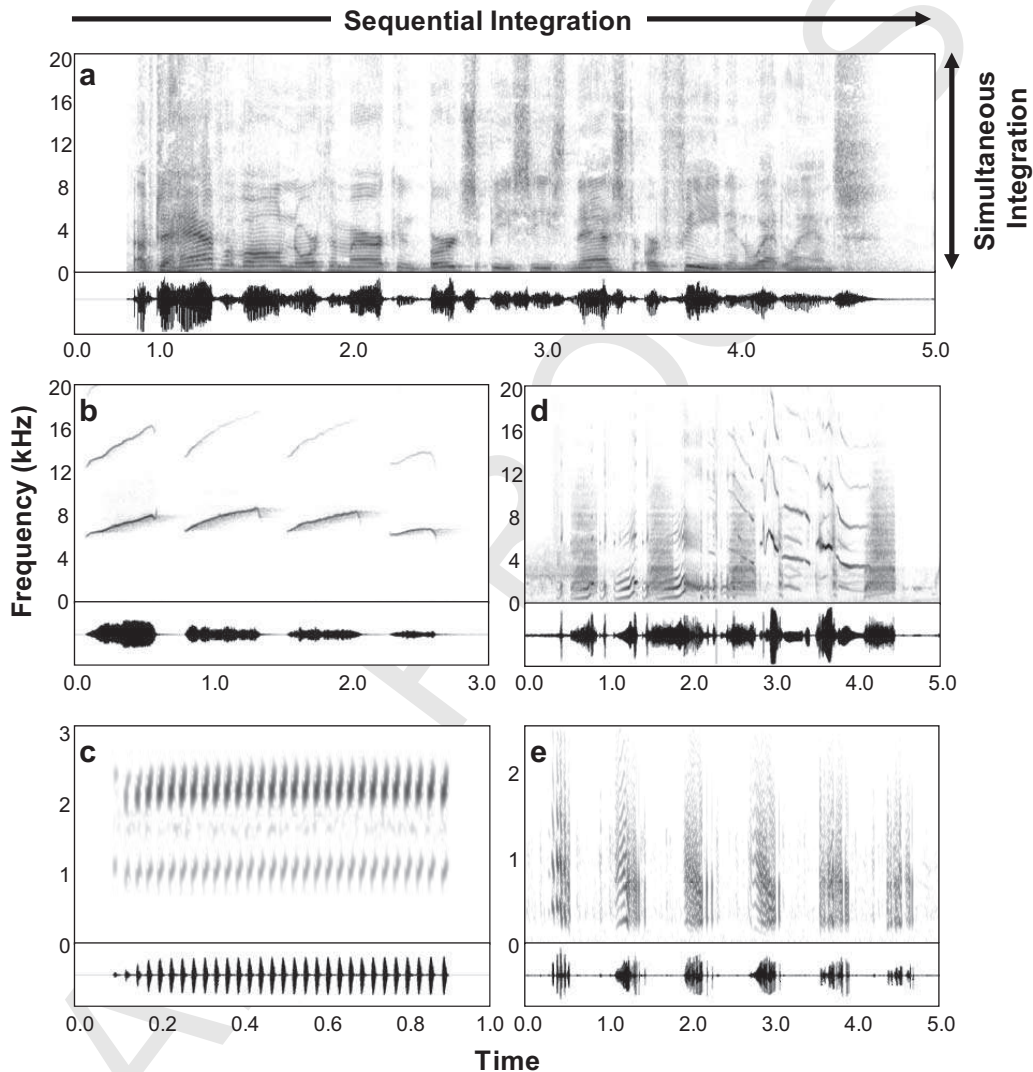

Figure 1. Spectrograms (top traces) and oscillograms (bottom traces) of animal vocalizations. A: Human speech ("Up to half of all North American bird species nest or feed in wetlands") spoken by President George W. Bush during an Earth Day celebration at the Laudholm Farm in Wells, Maine, on April 22, 2004 (courtesy of "The George W. Bush Public Domain Audio Archive" at http://thebots.net/GWBushSampleArchive.htm). B: "Phee" calls of the common marmoset, *Callithrix jacchus* (courtesy Rama Ratnam). C: Advertisement call of the gray treefrog, *Hyla chrysoscelis* (recorded by Mark Bee). D: Song motif from a European starling, *Sturnus vulgaris* (courtesy Lang Elliot). E: Portion of an advertisement call of the plains leopard frog, *Rana blairi* (recorded by Mark Bee). Note that in all cases, the vocalizations consist of sequences of sound elements (e.g., syllables and words [A], call notes [B,E], pulses [C], and song syllables [D]), many of which are comprised of simultaneous spectral components (e.g., harmonics), thus illustrating the potential necessity for sequential and simultaneous integration, as illustrated in Panel A.

The acoustic signals of nonhuman animals bear two general similarities with human speech that are relevant to our discussion. First, animal acoustic signals have gross temporal structure and often comprise sequences of sounds (Figure 1B–E). Second, many animal acoustic signals are harmonic (or quasi harmonic), meaning that the frequencies of concurrent spectral components are (approximately) integer multiples of the fundamental frequency (F0) (Figure 1B–E). Hence, in the context of animal acoustic communication, receivers may often face the two basic tasks of auditory scene analysis described previously: (a) the temporally separated sounds in a sequence of signal elements produced by the same individual must be integrated over time and segregated from the overlapping, interleaved, or alternating sound sequences from other signalers, and (b) simultaneous sounds (e.g., harmonics) that originate from the same individual must be perceptually grouped together and segregated from the concurrent sounds from other signalers.

Bottom-up and top-down processing. In humans, auditory scene analysis involves both bottom-up and top-down processes (reviewed in Bregman, 1990; Carlyon, 2004; Feng & Ratnam, 2000; Näätänen, Tervaniemi, Sussman, Paavilainen, & Winkler, 2001). Bottom-up mechanisms are *stimulus driven*, meaning that they operate only or primarily on cues present in the acoustic signal itself; they are largely automatic and obligatory, meaning that they do not critically depend on attention (although this is currently a disputed issue). In contrast, top-down processes depend on a listener's prior experience and expectations and thus involve higher level cognitive processes, such as learning, memory, and attention.

Many of the bottom-up mechanisms for auditory scene analysis probably operate at relatively low levels of the auditory system and may have arisen early in the evolution of vertebrate hearing (Fay & Popper, 2000; Feng & Ratnam, 2000; Hulse, 2002; Lewis & Fay, 2004; Popper & Fay, 1997). Spectral filtering and forward suppression are examples of neural processes that may mediate important bottom-up mechanisms in auditory scene analysis. These neural processes are observed in diverse species including monkeys (Fishman, Arezzo, & Steinschneider, 2004; Fishman, Reser, Arezzo, & Steinschneider, 2001), birds (Bee & Klump, 2004, 2005), and even insects (Schul & Sheridan, 2006), and they are already present at relatively low levels of the auditory system, such as the mammalian cochlear nucleus (Pressnitzer, Micheyl, Sayles, & Winter, 2007). These facts make it very likely that bottom-up auditory scene analysis functions in vocal communication across diverse taxa.

On the other hand, top-down processes appear to operate on the output of bottom-up processes occurring at lower levels of the auditory system. The extent to which top-down processes are involved in auditory scene analysis by nonhuman animals is an interesting and important question that has not been addressed. For example, compared to fish and frogs, we might expect the operation of top-down auditory scene analysis to be more prevalent in birds and mammals, which have relatively more complex auditory systems and for which various forms of vocal learning and vocally mediated social recognition can play important roles in acoustically mediated social behaviors (Hulse, 2002). However, the generality of such a statement could be questioned given that some fish (Myrberg & Riggio, 1985) and frogs (Bee & Gerhardt, 2002) also learn to recognize individuals by voice. In what follows, it

will be important to bear in mind that taxonomic differences among nonhuman animals could be reflected in potentially different contributions of top-down versus bottom-up processes to auditory scene analysis.

Acoustic cues for sequential and simultaneous integration. Due to the physics and biomechanics of sound production, the sounds produced by a given source or individual are more likely to share particular acoustic properties in common than are the sounds produced by different sources or individuals (Bregman, 1990, 1993; Cusack & Carlyon, 2004). Auditory systems appear to evolve to exploit these cues in the analysis of acoustic scenes. As described later, sounds that share common properties are more likely to be integrated together by the auditory system (i.e., commonalities promote integration). When properties differ enough between sound elements, they probably arose from different sources (or individuals), and these elements are more likely to be assigned to different auditory objects or streams (i.e., differences promote segregation). In humans, some of the acoustic properties of sound that play important roles in auditory scene analysis include fundamental frequency (F0) and harmonic relationships among spectral components (harmonicity), temporal onsets/ offsets, timbre, and patterns of amplitude modulation (reviewed in Bregman, 1990, 1993; Cusack & Carlyon, 2004; Darwin & Carlyon, 1995; Moore & Gockel, 2002). Much of what follows focuses on demonstrating how humans and other animals exploit cues related to these three acoustic properties in auditory scene analysis and in solving the cocktail party problem.

Integrating and Segregating Sequential Sounds: Auditory Streaming

In a cocktail-party-like environment, human listeners must perceptually segregate the sequences of speech sounds (e.g., syllables, words) spoken by different individuals. In other words, the auditory system must treat sounds emitted at different times by a given source as part of the same ongoing "stream," while at the same time separating those sounds from temporally adjacent sounds arising from other sources. Although these two processes involve stream integration and stream segregation, respectively, the processes that result in the formation of auditory streams are often referred to broadly as *auditory streaming* (Carlyon, 2004).

Auditory streaming can be easily demonstrated and studied in humans using stimulus sequences that consist of two tones of different frequencies, A and B, played in a repeating sequence, such as ABAB. . . or ABA–ABA–. . . (where – represents a silent gap; Figure 2A,B). The typical stimulus parameters of interest in **F2** such studies are the frequency separation (ΔF) between the A and B tones and the tone repetition time (TRT), which depends on tone duration and the intertone interval. When ΔF is relatively small (e.g., 1 semitone, or about 6%) and TRT relatively long (e.g., 5 tones/s), the percept is that of a single, coherent stream of one tone sequence that alternates in pitch (Figure 2A). In contrast, if ΔF is large (e.g., 10 semitones, or about 78%) and TRT is not too long, the percept becomes that of two separate streams corresponding to two separate sequences of A and B tones (Figure 2B). In this situation, the sensation of pitch alternation is lost, because the tones in each stream have a constant frequency and only one of the two streams can be attended at any time. These observations, which were initially reported by Miller and Heise (1950), have
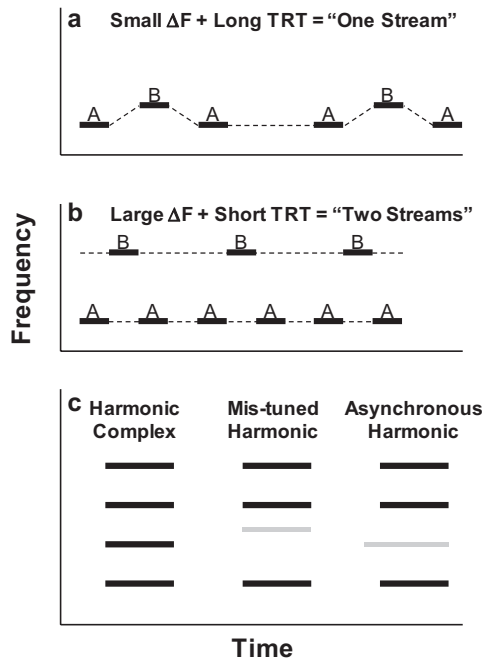
Figure 2. Schematic spectrograms illustrating experimental stimuli for investigating auditory streaming and simultaneous integration and segregation. A: An "ABA–ABA–. . ." tone sequence with a small difference in frequency (ΔF) between the A and B tones and a long tone repetition time (TRT); such a sequence would be perceived as a single, integrated stream of alternating tones with a galloping rhythm. B: An "ABA–ABA–. . ." tone sequence with a large ΔF and a short TRT; such a sequence would be perceived as two segregated streams, each with an isochronous rhythm. The dashed lines in Panels A and B indicate the percept (one vs. two streams, respectively). C: Three harmonic tone complexes showing a "normal" tone complex (left), a tone complex with a mistuned second harmonic (middle), and a tone complex with an asynchronous second harmonic that begins earlier than other the harmonics. In the later two cases, the second harmonic would likely be segregated from the rest of the integrated tone complex.

been confirmed and further investigated in many subsequent studies (reviewed in Bregman, 1990; Carlyon, 2004; Moore & Gockel, 2002).

Although highly stereotyped and simple, repeating two-tone sequences capture some essential features of the perceptual organization of more complex sound sequences. In particular, they illustrate the phenomenon of stream integration, whereby sounds are perceptually grouped or *bound* across time, yielding the percept of a coherent auditory stream that can be followed over time as a single entity. Two-tone sequences also serve to illustrate the converse phenomenon of stream segregation, whereby temporally proximal elements (e.g., the A and B tones in a sequence) are perceptually segregated, resulting in the perception of multiple streams of sounds that occur simultaneously and can be listened to selectively. In addition, with a proper choice of parameters, repeating tone sequences can be used to investigate top-down influences on the formation of auditory streams. Indeed, there is a relatively wide range of ΔFs and TRTs where the percept of one versus two streams depends on attentional factors (Carlyon, Cusack, Foxton, & Robertson, 2001; Carlyon, Place, Fantini, &

Cusack, 2003; Cusack, Deeks, Aikman, & Carlyon, 2004; Pressnitzer & Hupe, 2006; van Noorden, 1975). However, the influence of attention is also limited; under some stimulus conditions, the percept switches back and forth between one and two streams despite the listener's efforts to maintain a percept of either integrated or segregated streams (Pressnitzer & Hupe, 2006; van Noorden, 1975).

Auditory streaming and auditory detection, discrimination, and recognition. There is often a relationship between listeners' performance in auditory perception tasks involving sound sequences and the listeners' perception of those sound sequences as either integrated or segregated streams. For example, the identification of a temporal sequence of sounds, such as a melody, is facilitated by stimulus manipulations that promote its perceptual segregation from another, temporally interleaved sequence, such as presenting the melody to different ears, in a different frequency range, or with notes differing in timbre (Bey & McAdams, 2002, 2003; Cusack & Roberts, 2000; Dowling & Fujitani, 1971; Dowling, Lung, & Herrbold, 1987; Hartmann & Johnson, 1991; Iverson, 1995; Vliegen & Oxenham, 1999). Similar findings hold for interleaved sequences of synthetic vowels generated using different F0s or simulated vocal-tract sizes (Gaudrain, Grimault, Healy, & Bera, in press; Tsuzaki, Takeshima, Irino, & Patterson, 2007). In general, AQ: 4 the detection or discrimination of certain target sounds among other interfering sounds is facilitated under conditions that promote the perceptual segregation of targets from interferers, especially if targets and interferers share some subset of features in common that could otherwise cause them to be confused with each other (Gockel, Carlyon, & Micheyl, 1999; Micheyl & Carlyon, 1998; Micheyl, Carlyon, Cusack, & Moore, 2005) or when they vary rapidly and unpredictably over time (Kidd, Mason, & Arbogast, 2002; Kidd, Mason, & Dai, 1995; Kidd, Mason, Deliwala, Woods, & Colburn, 1994; Micheyl, Shamma, & Oxenham, 2007). AQ: 5 Conversely, there are some situations in which stream integration may be more advantageous than segregation. For instance, performance in the perception of the temporal order between consecutive sounds is usually higher when these sounds are perceived as part of a single stream (Bregman & Campbell, 1971; Brochard, Drake, Botte, & McAdams, 1999; Roberts, Glasberg, & Moore, 2002; Vliegen, Moore, & Oxenham, 1999).

Acoustic cues for auditory streaming. Sequences of alternating pure tones have helped to uncover some of the acoustic cues, such as ΔF and TRT, that determine whether successive sounds are likely to be integrated into a single stream or segregated into different streams (reviewed in Moore & Gockel, 2002). Recent neurophysiological studies have investigated these and other cues to discover where auditory streaming occurs in the brain (reviewed in Micheyl, Carlyon, et al., 2007). In highlighting the major role of tonotopic organization in promoting stream segregation, these neurophysiological studies are consistent with the so-called *channeling* theory of stream segregation (Beauvois & Meddis, 1996; Hartmann & Johnson, 1991). According to this theory, sounds that excite largely overlapping sets of peripheral filters (or *tonotopic channels*) tend to be heard as the same stream, whereas sounds that excite essentially nonoverlapping peripheral filters tend to be heard as different streams. Recent psychophysical studies, however, suggest that humans can even perceptually segregate sounds that excite the same peripheral channels into separate auditory streams based on difference in timbre and modulation rates (Gri-

mault, Bacon, & Micheyl, 2002; Grimault, Micheyl, Carlyon, Arthaud, & Collet, 2000; Vliegen et al., 1999; Vliegen & Oxenham, 1999). These findings have led to the suggestion that the auditory system takes advantage of any sufficiently salient perceptual difference between consecutive sounds to separate those sounds into different streams (Moore & Gockel, 2002).

Studies of auditory stream segregation in nonhuman animals. When sequences of repeated ABA– tone triplets are heard as a single stream, it evokes a distinctive galloping rhythm (Figure 2A); however, when the A and B tones perceptually split into two separate streams, this galloping rhythm is lost, and one hears two streams with isochronous tempi, one (A–A–A–A–. . .) three times faster than the other (–B——B——. . .; Figure 2B). Using the ABA– stimulus paradigm (Figure 2A, B) and operant conditioning techniques, MacDougall-Shackleton, Hulse, Gentner, and White (1998) took advantage of this perceived difference in rhythm between integrated and segregated streams to ask whether European starlings (*Sturnus vulgaris*) experienced stream segregation. After training starlings to discriminate between galloping and isochronous rhythms using single-frequency tone sequences, MacDougall-Shackleton et al. (1998) determined the probability that the birds would report hearing the repeating ABA– tones as a galloping sequence (one stream) or as isochronous sequences (two streams) as a function of increasing ΔF between the A and B tones. At large ΔFs, the birds more often reported hearing two streams. This result is important because it is consistent with observations in humans and provides strong evidence that at least one songbird also experiences the phenomenon of frequency-based stream segregation. Bee and Klump (2004, 2005) demonstrated neural correlates of these effects in starlings in a tonotopically organized area of the avian forebrain (field L2) that is the homologue of mammalian primary auditory cortex (for related work in macaque monkeys, see Fishman et al., 2001, 2004; Micheyl, Tian, Carlyon, & Rauschecker, 2005). Fay (1998, 2000) has reported generally similar findings on stream segregation using a classical conditioning paradigm with the goldfish (*Carassius auratus*).

Further indications that auditory streaming is experienced by nonhuman animals stems from experiments in which subjects were given a task in which performance was dependent on successful stream segregation. For example, Izumi (2001) measured the performance of Japanese macaques (*Macaca fuscata*) in the discrimination of short melodies in the absence or presence of interleaved *distractor* tones, which either did or did not overlap the melody in frequency. Note that this experiment is analogous to previous interleaved-melody recognition experiments in humans (Bey & McAdams, 2003; Dowling & Fujitani, 1971; Dowling et al., 1987; Vliegen & Oxenham, 1999). The results showed that the monkeys, like humans, were better at identifying the target melody when the distractor tones did not overlap spectrally with the target tones.

Auditory streaming and animal acoustic communication. Do animals that acoustically communicate in groups require the ability to perceptually integrate and segregate auditory streams? Here, we outline just two examples (out of many possible examples) for which we think auditory streaming could be at work. Consider first the case of song overlap in some songbirds (Naguib, 2005). During agonistic interactions, male songbirds can signal a high level of aggression or willingness to escalate the encounter by overlapping the songs of their opponents, whereas alternating songs indicates a relatively lower level of aggression (Dabelsteen, McGregor, Hol-

land, Tobias, & Pedersen, 1996, 1997; Naguib, 1999). Eavesdropping males and females that listen in on such sequences of song interactions in a communication network respond differently toward males that overlap their opponents' songs and males that either were overlapped or alternated with their opponent (Naguib, Fichtel, & Todt, 1999; Naguib & Todt, 1997; Otter et al., 1999; Peake, Terry, McGregor, & Dabelsteen, 2001, 2002). For eavesdroppers, determining which song elements were produced by different males in the interaction could be a case of auditory streaming in action: song elements from each bird would presumably have to be integrated together over time and segregated from the song elements of the other bird. The question of what sorts of spectral, temporal, and spatial (including distance) cues promote auditory streaming in eavesdroppers is an important one that has not been addressed.

A second example of acoustic signaling in groups that could involve auditory streaming involves signal timing interactions in frogs. These interactions range from near synchrony to complete alternation (see Figure 3), and they play important roles in female **F3** mate choice (Gerhardt & Huber, 2002; Grafe, 2005). In some species, such as the Kuvangu running frog (*Kassina kuvangensis*), two neighboring males can even precisely interdigitate the notes of
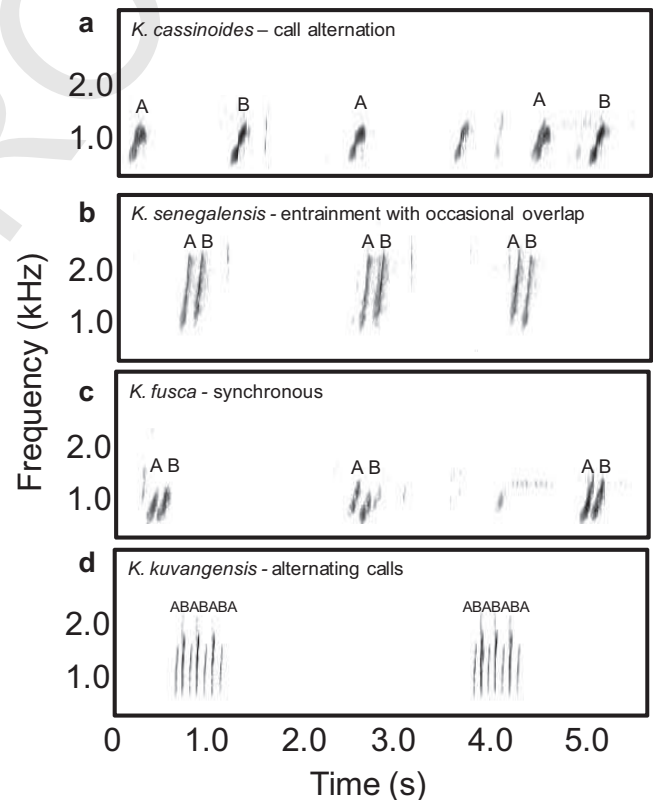
Figure 3. Spectrograms illustrating the diversity of call timing interactions in five species of frogs in the African genus *Kassina* (Grafe, 2005). **AQ: 21** A: Call alternation. B: Entrainment with occasional overlap. C: Synchrony. D: Entrainment with alternating calls. In each panel, the calls of two different males are labeled as "A" and "B." Note the general similarity between the alternating calls in Panel D and the artificial ABA– tone sequences depicted in Figure 2.

their calls (Figure 3D). Processes related to auditory streaming might be important in allowing females to choose one of the two signaling males. Note that the artificial ABA– and ABAB stimulus paradigms used to investigate auditory stream segregation in humans and starlings bear striking similarities to real-world problems for these particular frogs (cf. Figure 2 A, B and Figure 3). Could the cues for auditory streaming in humans and starlings identified using the ABA– paradigm also play a role in mate choice in frogs?

The important points to take from this section on the integration and segregation of sound sequences are the following: (a) diverse animal groups—for example, fish, songbirds, nonhuman primates—experience auditory streaming similar to that reported for humans using simple tone sequences; and (b) there are real-world examples of acoustic signaling interactions for which auditory streaming could be important. Thus far, however, few studies have investigated auditory streaming in the context of animal acoustic communication. One recent study of a katydid (*Neoconocephalus retusus*) suggested that spectral and temporal cues may allow them to segregate mating signals from the echolocation calls of predatory bats (Schul & Sheridan, 2006). An earlier study of the gray treefrog (*Hyla versicolor*) by Schwartz and Gerhardt (1995) suggested that a common spatial origin of sound elements comprising the pulsed advertisement call could contribute to call recognition via stream integration. Work by Farris et al. (Farris, Rand, & Ryan, 2002, 2005) on the túngara frog (*Physalaemus pustulosus*) suggested, however, that female frogs may in some cases be very permissive of large spatial separations when integrating temporally distinct signal elements across time.

We believe investigating the role of auditory streaming in animal acoustic communication is important, because streaming may directly relate to mechanisms that make possible certain functional behaviors in some animals, such as eavesdropping, mate choice, and predator detection. One important goal for future studies should be to ask whether and to what extent humans and other animals exploit the same acoustic cues in the formation of auditory streams and how mechanisms for exploiting these cues function in the acoustic signaling systems of nonhuman animals.

Integrating and Segregating Simultaneous Sounds

In a cocktail-party-like environment, human listeners not only must form auditory streams of speech that can be followed through time but also must perceptually integrate the simultaneous sounds originating from one person's voice (e.g., harmonics and speech formants) and segregate these from the concurrent sounds of other talkers. Here, we focus on three acoustic cues—harmonicity, onset synchrony, and common amplitude modulation—that produce potent perceptual effects in humans and are often regarded as the most important cues for promoting the integration and segregation of concurrent sounds. There is a strong intuitive appeal to assigning important roles to spatial information and source localization in perceptually segregating concurrent sounds. Studies of humans generally suggest, however, that spatial cues are probably relatively weak cues for the perceptual integration and segregation of concurrent sounds compared to other cues, such as harmonicity or onset synchrony (e.g., Culling & Summerfield, 1995; Darwin, 2006; Hukin & Darwin, 1995). In fact, there are many examples of the influence of spatial cues being overridden by other (acoustic) cues. A dramatic example of this is provided by Deutsch's octave

illusion (Deutsch, 1974), wherein sequences of alternating tones presented simultaneously to the two ears are organized perceptually based on frequency proximity rather than by ear of presentation. We return to the role of spatial cues in our discussion of spatial release from masking (see later discussion).

Harmonicity. Humans typically hear harmonic complex tones as a single fused sound with a unitary pitch corresponding to the F0, rather than as a series of separate pure tones with different pitches (Figure 2C, left). "Mistuning" one spectral component in an otherwise harmonic complex by more than about 2%–3%, however, causes it to "pop out," so that listeners hear two simultaneous sounds: a complex tone and a separate pure tone corresponding to the mistuned harmonic (Figure 2C, middle; Darwin, Ciocca, & Sandell, 1994; Hartmann, McAdams, & Smith, 1990; Moore, Glasberg, & Peters, 1986) . These two observations suggest that the auditory system takes advantage of harmonicity (i.e., common F0) for grouping together simultaneous spectral components that probably arose from the same source and for separating those components from inharmonically related components, which probably arose from a different source. Although this conclusion has been qualified by Roberts and colleagues (Roberts & Bailey, 1996a, 1996b; Roberts & Brunstrom, 1998, 2001), who have shown in a series of elegant studies that the perceptual fusion of spectral components depends on regular spectral spacing rather than harmonicity per se, most natural sounds with evenly spaced spectral components are likely to be also harmonic or quasi harmonic.

In humans, differences in F0, and thus inharmonic relationships among spectral components, usually result in substantial improvements in the identification of concurrent speech sounds, ranging from whole sentences (Brokx & Nooteboom, 1982) to isolated vowels (Culling & Darwin, 1993; de Cheveigne, McAdams, Laroche, & Rosenberg, 1995; Scheffers, 1983; Summerfield & Assmann, 1991; Zwicker, 1984). In the latter case, the effect is unlikely to be mediated by perceived segregation because it occurs over a range of F0 separations that are too small to evoke a percept of two separate sources (Assmann & Summerfield, 1994; Culling & Darwin, 1994). Another line of evidence for the beneficial influence of F0 differences on the perceptual segregation of concurrent sounds comes from findings that listeners can more easily identify and finely discriminate the pitch of a target harmonic complex mixed together with another complex (or *interferer*), if the two complexes have different F0s than if their F0s are similar. These benefits of differences in F0, which hold whether the target and interferer occupy the same (Beerends & Houtsma, 1986, 1989; Carlyon, 1996; Micheyl, Bernstein, & Oxenham, 2006) or distinct (Gockel, Carlyon, & Plack, 2004; Micheyl & Oxenham, 2007) spectral regions, probably contribute to why it is intuitively (and objectively) easier to follow a female speaker in the presence of a male interferer (or vice versa) than when the target and interferer are both of the same gender and thus more likely to have similar F0s. Thus, overall, many results in the human psychophysical literature indicate that differences in F0 provide powerful cues for the perceptual separation of concurrent harmonic sounds, whereas a common F0 and harmonic relationships (or common spectral spacing) among spectral components promote perceptual fusion (reviewed in Darwin & Carlyon, 1995).

Onset synchrony. Frequency components that start and end at the same time tend to be perceptually grouped together, whereas

components that start at (sufficiently) different times tend to be heard as separate sounds (Figure 2C, right; reviewed in Darwin & Carlyon, 1995). For instance, Bregman and Pinker (1978) showed that listeners tended to "hear out" the two spectral components of a two-tone complex as separate tones when they were asynchronous. More objective evidence that onset asynchronies can greatly facilitate the selective extraction of information from a target sound in the presence of another sound comes from findings that listeners can more accurately perceive (i.e., identify or discriminate) the pitch of a complex tone (target) in the presence of another complex tone (interferer), if the target tone starts later and ends earlier than the interferer. This effect has been observed when the target and interferer sounds occupied either the same spectral region (Micheyl et al., 2006) or nonoverlapping spectral regions (Gockel et al., 2004; Micheyl & Oxenham, 2007).

The influence of onset synchrony as an auditory grouping cue in speech perception was demonstrated in several elegant studies by Darwin and colleagues (Ciocca & Darwin, 1993; Darwin, 1984; Darwin & Ciocca, 1992; Darwin & Hukin, 1998; Darwin & Sutherland, 1984; Hill & Darwin, 1996; Hukin & Darwin, 1995). In particular, these authors showed that the phonemic identity of a synthetic vowel could be altered by making one of the harmonics close to a formant peak start earlier than the others. This result was interpreted as indicating that an asynchronous onset promoted the perceptual segregation of the temporally shifted component from the rest of the vowel, resulting in a shift of the perceived frequency of the formant peak closest to the perceptually removed component.

Common amplitude modulation. Sounds in the real world are often broadband and fluctuate in level, that is, they are amplitude modulated (Richards & Wiley, 1980; Singh & Theunissen, 2003). In addition, these amplitude modulations may often be correlated in time across different regions of the frequency spectrum (Klump, 1996; Nelken, Rotman, & Bar Yosef, 1999). Studies of two phenomena known as *comodulation masking release* (CMR; Hall, Haggard, & Fernandes, 1984) and *comodulation detection difference* (CDD; McFadden, 1987) indicate that the human auditory system is able to exploit correlated envelope fluctuations across the frequency spectrum (i.e., comodulation) to improve the detection of signals presented with concurrent masking noise (reviewed **AQ: 6** in Hall, Grose, & Mendoza, 1995; Langemann & Klump,; Verhey, Pressnitzer, & Winter, 2003).

Two experimental paradigms have been used to investigate CMR in humans (Verhey et al., 2003). In the *band-widening* **F4** *paradigm* (Figure 4A), the bandwidth of a bandpass noise centered on the frequency of the target tone is varied between a narrow bandwidth (e.g., within a single auditory filter) and much wider bandwidths (e.g., spanning multiple auditory filters). At narrow bandwidths, signal detection thresholds are generally similar in the presence of comodulated and unmodulated maskers. As the bandwidth is increased beyond the critical bandwidth of the auditory filter centered on the target tone, however, signal detection thresholds become lower (i.e., signal detection is easier) in the presence of comodulated noise (Figure 4A) but remain fairly constant in unmodulated noise, even though the overall levels of the two types of maskers are the same. Hence, in the band-widening paradigm, the benefits in signal detection in comodulated noise are most pronounced when the bandwidth of the masker is sufficiently wide to span multiple auditory filters (reviewed in Verhey et al., 2003).
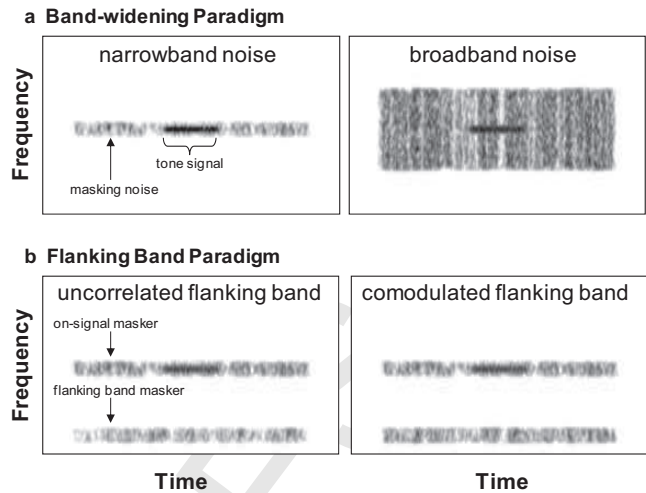

Figure 4. Schematic spectrograms illustrating experimental paradigms for investigating comodulation masking release (CMR) for detecting a short tone signal. A: The band-widening paradigm showing a modulated narrowband noise (left) and a modulated broadband noise (right). B: The flanking band paradigm showing the on-signal masker, the flanking band masker, and either the uncorrelated condition (left) or the comodulated condition (right). In the schematic examples depicted here, the magnitude of CMR would be greater in the conditions illustrated in the right panel for both paradigms.

A second approach to investigate CMR—the *flanking band paradigm* (Figure 4B)—uses a masker comprising a narrowband noise centered on the frequency of the target tone (the *on-signal band*) and one or more narrowband noises (*flanking bands*) spectrally located at frequencies remote from that of the target signal and on-signal band. Typically, flanking bands have center frequencies that fall outside of the critical bandwidth of the auditory filter centered on the target signal. In this paradigm, signal detection thresholds are typically lower when the on-signal band and the flanking bands have comodulated envelopes compared to conditions in which their envelopes fluctuate independently (reviewed in Verhey et al., 2003). Studies of CDD use a generally similar flanking band paradigm with two exceptions: (a) the target signal is a modulated narrow band of noise and not a tone, and (b) there is no on-signal masking band. In studies of CDD (reviewed in Langemann & Klump,), thresholds for detecting the narrowband **AQ: 7** noise signal are lower when the flanking bands all share a common envelope that is different from that of the signal. Thresholds are higher either when the signal envelope is correlated with the comodulated flanking bands or when the signal and all flanking bands have independently fluctuating envelopes. What these studies of CMR and CDD demonstrate is that the human auditory system is sensitive to temporal correlations in amplitude fluctuation across the frequency spectrum.

One hypothesis for improved signal detection when signal envelopes fluctuate independently of a comodulated masker is that common amplitude modulations across the spectrum of the masker promote the across-channel integration of masker energy into an auditory object that is distinct from the independently fluctuating signal. Although there is evidence to support this hypothesis for CMR (Verhey et al., 2003), within-channel processes unrelated to

auditory grouping may be more important in CDD (Buschermöhle, Feudel, Klump, Bee, & Freund, 2006). Moreover, the extent to which processes related to CMR and CDD play a role in speech perception in cocktail-party-like environments is not yet clear. Early studies suggested that CMR-like processes might play a role in speech detection but probably contributed little to speech intelligibility at suprathreshold levels (Festen, 1993; Grose & Hall, 1992). More recent work, however, suggests that CMR-like processes might also contribute to speech recognition (Kwon, 2002). The debates about whether CMR and CDD share common underlying across-channel mechanisms and whether these phenomena contribute more to signal detection than signal recognition go beyond this review. The important question for our purposes is whether processes related to CMR and CDD might operate similarly in humans and nonhuman animals and contribute to acoustic signal perception in nonhuman animals.

Studies of integrating and segregating simultaneous sounds in nonhuman animals. Only a few studies have specifically addressed whether the acoustic cues that promote auditory grouping and segregation of concurrent sounds in humans also function in the acoustic communication systems of nonhuman animals. Playback studies with bullfrogs (*Rana catesbeiana*; Simmons & Bean, 2000) and cotton-top tamarins (*Saguinus oedipus*; Weiss & Hauser, 2002), for example, suggest that harmonicity could be an important cue in the perception of vocal communication signals in these species. Geissler and Ehret (2002) demonstrated in mice (*Mus musculus*) that onset synchrony between the harmonics in pup wriggling calls were important in allowing mothers to form coherent auditory objects of these vocalizations.

To our knowledge, no study of nonhuman vertebrates has directly investigated CMR or CDD in the context of how receivers perceive acoustic communication signals. Traditional psychophysical experiments, however, have revealed that at least one well-studied songbird experiences magnitudes of CMR and CDD similar to those reported for humans. Klump and Langemann (1995) used operant conditioning techniques and the band-widening paradigm to show that starlings (*S. vulgaris*) experience a mean CMR of about 11 dB in a tone detection task, which was similar to that reported in a similar study of humans (e.g., Schooneveldt & Moore, 1989). Klump and Nieder (2001) later reported neural correlates of these findings based on recordings from the starling forebrain (field L2). As in humans, starlings also experienced CMR in a flanking band paradigm (Langemann & Klump, 2001), and neural correlates of CMR using this paradigm have also been found in the responses of starling forebrain neurons (Hofer & Klump, 2003; Nieder & Klump, 2001). Starlings also experience a magnitude of CDD similar to that found in humans (Langemann & Klump,), and, not surprisingly perhaps, correlates of CDD can be found in the responses of starling forebrain neurons (Bee, Buschermöhle, & Klump, ; Buschermöhle et al., 2006). A CDD effect similar to that observed in humans has also been reported in a recent study of the hooded crow (*Corvus corone cornix*; Jensen, 2007).

Studies of CMR in nonhuman animals have not been limited to birds. At a behavioral level, CMR has also been demonstrated in the Mongolian gerbil (*Meriones unguiculatus*; Klump, Kittel, & Wagner, 2001). Improvements in signal detection related to CMR have been demonstrated at a neurophysiological level in the auditory systems of leopard frogs (*R. pipiens*; Goense, 2004), guinea

pigs (*Cavia porcellus*; Neuert, Verhey, & Winter, 2004), and cats (*Felis catus*; Nelken, Jacobson, Ahdut, & Ulanovsky, 2000; Nelken et al., 1999). Together, these studies of CMR and related phenomena suggest that nonhuman vertebrates across a range of taxa possess neural mechanisms that could function to exploit common amplitude modulations across the frequency spectrum of masking noises to improve signal perception. We are aware of only one study that has investigated CMR-related effects in the context of acoustic communication in invertebrates. In that study, Ronacher and Hoffmann (2003) found little evidence for the operation of CMR-like processes in the phonotaxis response of male grasshoppers (*Chorthippus biguttulus*) to female stridulation signals. Could this represent a difference in signal-processing strategies between insects and vertebrates?

Additional studies of the roles of harmonicity and onset synchrony in the acoustic communication systems of nonhuman animals would make valuable contributions to our understanding of the extent to which these acoustic cues function in the integration and segregation of the concurrent sound elements comprising acoustic communication signals (Figure 1B–E). Future studies that quantify the modulation statistics of the natural acoustic scenes in which animals communicate (e.g., Nelken et al., 1999; Singh & Theunissen, 2003) and test the hypothesis that common amplitude modulations could be exploited to improve signal detection and recognition would also make valuable contributions to our understanding of the mechanisms of acoustic signal perception in nonhuman animals.

Spatial Release From Masking

Intuitively, it would seem that the integration and segregation of both sequential and simultaneous sounds could be as easy as assigning interleaved or concurrent sounds to the different locations in space corresponding to the positions of different sound sources. It is surprising, however, that many human psychoacoustic studies have now shown that spatial cues related to interaural time differences (ITDs) or interaural level differences (ILDs) play a relatively weak role in perceptual integration and segregation compared to other cues (e.g., Culling & Summerfield, 1995; Darwin, 2006; Hukin & Darwin, 1995). This is not to say, however, that spatial cues play no role in solving the cocktail party problem. In humans, speech intelligibility under cocktail-party-like listening conditions is improved when there is spatial separation between a source of target speech and interfering sources of speech or speechlike masking noise (Bronkhorst, 2000; Freyman, Balakrishnan, & Helfer, 2001; Hawley, Litovsky, & Culling, 2004; Shinn-Cunningham et al., 2005; Shinn-Cunningham, Schickler, Kopco, & Litovsky, 2001). For example, compared to conditions in which sources of target speech and interfering speech or noise with the spectrum and envelope modulations of speech are presented from the same frontal direction, listeners experience a 6–10 dB "release" from masking when the masker is displaced 90° lateral to the signal in speech recognition tasks (Bronkhorst, 2000). This general phenomenon is known as *spatial release from masking* or *spatial unmasking*.

Spatial release from masking of speech can result from three causes under binaural listening conditions (Bronkhorst, 2000; Shinn-Cunningham et al., 2005). First, when the masker is displaced in azimuth relative to a frontally presented signal, the head

creates a sound shadow that attenuates the level of the masker at one ear (the so-called *best ear for listening*) and results in ILDs for the masker but not the signal. Second, the displacement of the masker to one side of a frontally presented signal creates ITDs in the arrival of the masker, but not the signal, at the two ears. Finally, the locations from which the listener actually perceives the signal and masker as originating can influence the magnitude of spatial unmasking through processes thought to be related to spatial attention (Freyman et al., 2001; Shinn-Cunningham et al., 2005). For our purposes, it is important to note that the magnitudes of ITDs and ILDs vary directly with head size. Given that humans are notable for their large heads compared to most animals, should we expect nonhuman animals to experience magnitudes of spatial release from masking similar to those experienced by humans? Current evidence suggests that this is often the case.

Studies of spatial unmasking in nonhuman animals. A few studies have used traditional psychophysical techniques to investigate spatial unmasking in animals. In a study of budgerigars (*Melopsittacus undulates*), Dent et al. (Dent, Larsen, & Dooling, 1997) reported a maximum spatial release from masking of about 10 dB when a target tone and masking noise were separated by 90°. Similar findings have also been reported for ferrets (*Mustela putorius*; Hine, Martin, & Moore, 1994). Holt and Schusterman (2007) recently reported results from a study of spatial unmasking of airborne sounds in the harbor seal (*Phoca vitulina*) and the California sea lion (*Zalophus californianus*). Depending on the frequency of the tone signal, the magnitude of spatial unmasking when an octave-band noise centered on the tone frequency was separated by 90° ranged from 8.5 to 19.0 dB and −1.3 to 11.7 dB in the harbor seal and sea lion, respectively. The magnitude of spatial release from masking in harbor seals for detecting airborne sounds was slightly larger than that previously reported in a study of underwater spatial unmasking in this species (Turnbull, 1994). Ison and Agrawal (1998) used a reflex modification technique to demonstrate spatial release from masking in the mouse (*M. musculus*) but did not report the magnitude of masking release (in dB). One particularly interesting recent study demonstrated that humans experience spatial release from masking in a task in which listeners were required to identify zebra finches (*Taeniopygia guttata*) by their individually distinct songs in the presence of an artificial zebra finch chorus (Best, Ozmeral, Gallun, Sen, & Shinn-Cunningham, 2005). It would be interesting to know how the performance of zebra finches compares to that of humans in the same task. More generally, it will be important to investigate in future studies how the performance of animals in relatively simple tone detection tasks compares with their performance in spatial unmasking tasks that require the detection and recognition of conspecific communication signals.

Two studies have used phonotaxis assays to estimate the magnitude of spatial unmasking in the context of acoustic communication in frogs. In a study by Schwartz and Gerhardt (1989) of green treefrogs (*H. cinerea*), females were required to behaviorally discriminate between attractive advertisement calls and less attractive aggressive calls separated by 180°° around the perimeter of a circular test arena. In separate conditions, the positions of two broadband (0.1–10 kHz) noise sources were such so that either one noise source was located next to each signal speaker or each noise source was located 45° from each signal speaker or 90° lateral to both signal speakers and on opposite sides of the arena (180°

apart). The maximum magnitude of spatial release from masking reported in this study was about 3 dB for signal detection, but there was little evidence to suggest a spatial release from masking in call discrimination. More recently, Bee (in press) reported a 6–12 dB release from masking in the gray treefrog (*H. chrysoscelis*) when a steady-state masker with the spectrum of a natural breeding chorus was located 90° lateral to a speaker broadcasting target advertisement calls compared to a condition in which the signal and chorus-shaped noise were separated by only 7.5°. These results for gray treefrogs are similar to what might have been predicted based on neurophysiological studies of the northern leopard frog (*R. pipiens*) in which the magnitude of spatial release from masking observed in neural signal detection thresholds was about 9 dB in the inferior colliculus (torus semicircularis), which receives binaural inputs (Lin & Feng, 2003; Ratnam & Feng, 1998).

Together, these studies of animals suggest that spatial release from masking is not at all unique to humans. The behavioral and neurophysiological studies of frogs indicate that spatial release from masking could function in the context of acoustic communication and, thus, could be one important process that allows some nonhuman animals to cope with cocktail-party-like problems. For frogs (Gerhardt & Bee, 2006; Gerhardt & Huber, 2002) and perhaps birds (Klump & Larsen, 1992; Larsen, Dooling, & Michelsen, 2006), the internal coupling of the two inner ears and the operation of pressure-difference systems probably compensate for some of the limitations on directional hearing that would otherwise result from a small head size. Ronacher and Hoffmann (2003) found little evidence for spatial release from masking in a study of the grasshopper (*C. biguttulus*). Hence, there are potentially interesting taxonomic differences in the operation of spatial release from masking in the context of acoustic communication.

It is also worth making explicit here that an approaching using one or a limited number of masking noise sources in highly controlled laboratory studies of spatial unmasking does not wholly reflect the real-world listening conditions that many animals face. Nevertheless, such an approach almost certainly engages perceptual mechanisms that are important for solving cocktail-party-like problems in nature. Moreover, this approach constitutes an important and often necessary first step toward understanding the role of spatial unmasking in solving the cocktail party problem.

Why Should Animal Behaviorists Study the Cocktail Party Problem?

As the studies cited in the previous sections indicate, humans and some nonhuman animals probably face similar problems and use similar solutions when it comes to perceiving acoustic signals in noisy social environments comprised of groups of simultaneously signaling individuals. Given the interest among animal behaviorists in acoustic signaling interactions in groups, studies of auditory scene analysis and the cocktail party problem have probably received less attention than is warranted (Hulse, 2002). We believe there are excellent reasons why animal behaviorists should study the cocktail party problem.

Evolutionary Diversity in Sensory Mechanisms

Understanding the mechanistic causes of behavior has long been an important goal of animal behavior research (Tinbergen, 1963).

Indeed, studies of the mechanisms of signal production and perception have a long and rich history in the study of animal communication (Bradbury & Vehrencamp, 1998). The mechanisms of auditory scene analysis, the sensory solutions to cocktail-party-like problems, and their role in animal acoustic communication represent opportunities for future research that will almost certainly yield new and important insights into the mechanisms and evolution of both hearing and acoustic communication. The sense of hearing arose early in the evolution of the vertebrates. Consequently, some of the basic processing strategies involved in auditory scene analysis may be shared (i.e., synapomophic) between humans and other nonhuman vertebrates (Popper & Fay, 1997). However, many different taxonomic groups may have independently evolved communication systems that involve acoustic signaling in groups or large aggregations (e.g., insects and frogs and songbirds). Within the constraints imposed by the physics of sound, natural selection may have elaborated on certain basic mechanisms for auditory scene analysis in different taxonomic groups to produce a diversity of evolutionarily derived (i.e., apomorphic) sensory solutions to cocktail-party-like problems. These novel mechanisms often cannot be predicted but must be uncovered through empirical studies. Thus, one potential contribution of future research on animal cocktail parties could be the discovery of both shared and derived sensory mechanisms underlying the perception of acoustic signals in noisy social settings.

Receiver Psychology and Communication Network Theory

Arguably two of the most important conceptual advances in the recent study of animal communication are the ideas of *receiver psychology* and *communication network theory*. Our understanding of both would benefit from a better understanding of the role of auditory scene analysis in animal acoustic communication.

Receiver psychology holds that the evolution of both signal design and the behavioral strategies that signalers use depend critically on the processing strategies of a receiver's nervous system (Guilford & Dawkins, 1991; Rowe, 1999; Rowe & Skelhorn, 2004). Precisely how the physiological mechanisms underlying auditory scene analysis might have influenced the evolution of signals and signaling behaviors remains an important but still largely unexplored question. We also should not be surprised if under some conditions (e.g., extremely dense aggregations) there are evolutionary constraints on the extent to which signals and signaling strategies can be modified to improve signal perception by receivers (Bee, in press). Thus, for animals that communicate acoustically in groups, the sensory solutions to the cocktail party problem and the broader concept of auditory scene analysis deserve consideration in the context of receiver psychology (Bee, in press).

Among the most important contributions of communication network theory (McGregor, 2005) is the hypothesis that animals can gather information by eavesdropping on the signaling interactions that occur among two or more signalers (Peake, 2005). There is now considerable evidence in support of this hypothesis, especially among songbirds (reviewed in Peake, 2005). Most previous studies of eavesdropping have emphasized the amount and adaptive value of information gathered by eavesdroppers, but eavesdropping also represents a quite interesting problem when the perceptual mechanisms that make it possible are considered. As illustrated earlier, eavesdropping in acoustic communication networks would seem to demand capabilities of auditory scene analysis (Langemann & Klump, 2005). Our understanding of the perceptual mechanisms that make possible the extraction of information in acoustic communication networks would benefit by approaching these issues from an auditory scene analysis perceptive.

Anthropogenic Noise

Anthropogenic noise represents an evolutionarily recent intruder into the acoustic scenes that humans and other animals have evolved to analyze. There is increasing concern among animal behaviorists and conservation biologists that noise pollution could interfere with animal acoustic communication systems (Katti & Warren, 2004; Patricelli & Blickley, 2006; Rabin & Greene, 2002; Slabbekoorn, Yeh, & Hunt, 2007; Warren, Katti, Ermann, & Brazel, 2006). Recent experimental evidence from frogs suggests that anthropogenic noise can inhibit calling by males (Sun & Narins, 2005) and mask the perception of signals by females (Bee & Swanson, in press; see Lohr et al., 2003, for related work in birds). Some animals, such as songbirds and cetaceans, may be able to modify their signals in ways that ameliorate the effects of high anthropogenic noise levels (Fernández-Juricic et al., 2005; Foote, Osborne, & Hoelzel, 2004; Slabbekoorn & den Boer-Visser, 2006; Slabbekoorn & Peet, 2003; Wood & Yezerinac, 2006), but this solution may not work for all animals (Bee & Swanson, in press). Moreover, as illustrated previously, animals also possess mechanisms for coping with masking noise (e.g., CMR and spatial release from masking). A better understanding of auditory scene analysis, and the general mechanisms that operate to segregate behaviorally relevant signals from noise, will be necessary to accurately assess the magnitude of the threat that anthropogenic noise poses to animal acoustic communication systems.

Applications to Humans

Behavioral and physiological studies of auditory scene analysis in a diversity of acoustically communicating animals could become relevant to translational research on auditory signal processing by humans and machines. Much of the interest in human auditory scene analysis and the cocktail party problem stems from the fact that people with hearing impairments have much greater difficulty understanding speech under noisy conditions when compared to listeners with healthy auditory systems. For example, compared to listeners with normal hearing, listeners with hearing impairments have difficulty in auditory stream segregation tasks with complex tones (e.g., Grimault, 2004; e.g., Grimault, Micheyl, Carlyon, Arthaud, & Collet, 2001) and experience less comodulation masking release (e.g., Moore, Shailer, Hall, & Schooneveldt, 1993) and less spatial release from masking (e.g., Bronkhorst, 2000; Bronkhorst & Plomp, 1992). Whereas hearing aids and cochlear implants improve speech perception in quiet settings, they typically provide their users with much less benefit in noisy, real-world situations (Moore, Peters, & Stone, 1999; Stickney, Zeng, Litovsky, & Assmann, 2004). Basic research on both the physiological mechanisms of auditory scene analysis in a diversity of animal species and their function in animal acoustic communi-

cation systems might ultimately contribute to improvements in hearing prosthetic technology. For example, findings from studies of directional hearing in the fly *Ormia ochracea*, an acoustically orienting parasitoid of signaling crickets, have already had implications for the development of miniscule directional microphones for hearing aids (Mason, Oshinsky, & Hoy, 2001; Miles & Hoy, 2006).

In the field of computational auditory scene analysis, efforts to apply the principles of human auditory scene analysis to computer algorithms for automated speech recognition have met with some success (Cooke & Ellis, 2001; Wang & Brown, 2006). Compared to relatively quiet conditions, however, artificial speech recognition systems exhibit drastically reduced performance in noisy situations with competing speech signals (Barker, 2006; Lippmann, 1997). A better understanding of auditory scene analysis in nonhuman animals might ultimately broaden and deepen our understanding of the potential diversity of the physiological mechanisms whereby auditory scene analysis is accomplished and thereby contribute to the design of biologically inspired artificial scene analysis systems.

General Conclusions

Animals that acoustically signal in groups or large social aggregations often encounter and solve problems closely related to the human cocktail party problem. All of the mechanisms discussed previously that facilitate human speech perception in noisy social situations can also be identified as auditory processes operating in nonhuman animals. Few studies, however, have explicitly investigated these mechanisms in the context of animal acoustic communication. As Hulse (2002) recently lamented in his review of this topic, "sometimes scene analysis is so obvious it is overlooked." To be sure, other mechanisms that we have not touched upon may also play important roles in allowing animals to solve cocktail-party-like problems. For example, in animals with multimodal signals, the cross-modal integration of acoustic signals with visual signals or other cues associated with the acoustic signals might improve acoustic signal perception in cocktail-party-like social environments (e.g., Ghazanfar & Logothetis, 2003; Partan & Marler, 1999; Rowe, 2002). Answering questions about how animals solve cocktail-party-like problems will require that future studies adopt innovative approaches that integrate questions, methods, and hypotheses from previous psychoacoustic studies of humans with behavioral and neuroethological studies of acoustic communication in a wide diversity of animals. We also wish to stress that the use of seemingly artificial or highly contrived experimental stimuli under controlled laboratory conditions may often be the best way, and perhaps sometimes the only way, to investigate the sensory solutions to cocktail-party-like problems at deeper levels than would be possible using strictly natural stimuli presented in the animals' natural habitats. Even though approaches to studying the cocktail party problem in animals will certainly pose new challenges, the benefits of adopting the auditory scene analysis paradigm would significantly advance the study of animal acoustic communication in much the same way that it has already contributed to our understanding of human hearing and speech communication.

AQ: 10

References

Assmann, P. F., & Summerfield, Q. (1994). The contribution of waveform interactions to the perception of concurrent vowels. *Journal of the Acoustical Society of America, 95,* 471–484.

Aubin, T., & Jouventin, P. (1998). Cocktail-party effect in king penguin colonies. *Proceedings of the Royal Society of London Series B-Biological Sciences, 265,* 1665–1673.

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.

Barker, J. (2006). Robust automatic speech recognition. In D. Wang & G. J. Brown (Eds.), *Computational auditory scene analysis: Principles, algorithms, and applications* (pp. 297–350). Hoboken, NJ: Wiley-IEEE Press.

Beauvois, M. W., & Meddis, R. (1996). Computer simulation of auditory stream segregation in alternating-tone sequences. *Journal of the Acoustical Society of America, 99,* 2270–2280.

Bee, M. A. (in press). Sound source segregation in grey treefrogs: Spatial release from masking by the sound of a chorus. *Animal Behaviour.*

Bee, M. A., Buschermöhle, M., & Klump, G. M. (). Detecting modulated signals in modulated noise: II. Neural thresholds in the songbird forebrain. Manuscript submitted for publication.   **AQ: 11**

Bee, M. A., & Gerhardt, H. C. (2002). Individual voice recognition in a territorial frog (*Rana catesbeiana*). *Proceedings of the Royal Society of London Series B-Biological Sciences, 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,* 197–214.

Bee, M. A., & Swanson, E. M. (in press). Auditory masking of anuran advertisement calls by road traffic noise. *Animal Behaviour.*

Beecher, M. D. (1989). Evolution of parent-offspring recognition in swallows. In D. A. Dewsbury (Ed.), *Contemporary issues in comparative psychology* (pp. 360–380). Sunderland, MA: Sinauer.

Beecher, M. D. (1991). Successes and failures of parent-offspring recognition systems in animals. In P. G. Hepper (Ed.), *Kin recognition* (pp. 94–124). Cambridge, UK: Cambridge University Press.

Beerends, J. G., & Houtsma, A. J. (1986). Pitch identification of simultaneous dichotic two-tone complexes. *Journal of the Acoustical Society of America, 80,* 1048–1056.

Beerends, J. G., & Houtsma, A. J. (1989). Pitch identification of simultaneous diotic and dichotic two-tone complexes. *Journal of the Acoustical Society of America, 85,* 813–819.

Benney, K. S., & Braaten, R. F. (2000). Auditory scene analysis in estrildid finches (*Taeniopygia guttata* and *Lonchura striata domestica*): A species advantage for detection of conspecific song. *Journal of Comparative Psychology, 114,* 174–182.

Best, V., Ozmeral, E., Gallun, F. J., Sen, K., & Shinn-Cunningham, B. G. (2005). Spatial unmasking of birdsong in human listeners: Energetic and informational factors. *Journal of the Acoustical Society of America, 118,* 3766–3773.

Bey, C., & McAdams, S. (2002). Schema-based processing in auditory scene analysis. *Perception & Psychophysics, 64,* 844–854.

Bey, C., & McAdams, S. (2003). Postrecognition of interleaved melodies as an indirect measure of auditory stream formation. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 267–279.

Bradbury, J. W., & Vehrencamp, S. L. (1998). *Principles of animal communication.* Sunderland, MA: Sinauer.

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound.* Cambridge, MA: MIT Press.

Bregman, A. S. (1993). Auditory scene analysis: Hearing in complex environments. In S. McAdams & E. Bigand (Eds.), *Thinking in sound:*

The cognitive psychology of human audition (pp. 10–36). Oxford: Clarendon Press.

Bregman, A. S., & Campbell, J. (1971). Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of Experimental Psychology, 89,* 244–249.

Bregman, A. S., & Pinker, S. (1978). Auditory streaming and building of timber. *Canadian Journal of Psychology, 32,* 19–31.

Brochard, R., Drake, C., Botte, M. C., & McAdams, S. (1999). Perceptual organization of complex auditory sequences: Effect of number of simultaneous subsequences and frequency separation. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 1742–1759.

Brokx, J. P. L., & Nooteboom, 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.

Bronkhorst, A. W., & Plomp, R. (1992). Effect of multiple speech-like maskers on binaural speech recognition in normal and impaired hearing. *Journal of the Acoustical Society of America, 92,* 3132–3139.

Brumm, H., & Slabbekoorn, H. (2005). Acoustic communication in noise. *Advances in the Study of Behavior, 35,* 151–209.

Buschermöhle, M., Feudel, U., Klump, G. M., Bee, M. A., & Freund, J. (2006). Signal detection enhanced by comodulated noise. *Fluctuation and Noise Letters, 6,* 339–347.

Carlyon, R. P. (1996). Encoding the fundamental frequency of a complex tone in the presence of a spectrally overlapping masker. *Journal of the Acoustical Society of America, 99,* 517–524.

Carlyon, R. P. (2004). How the brain separates sounds. *Trends in Cognitive Sciences, 8,* 465–471.

Carlyon, R. P., Cusack, R., Foxton, J. M., & Robertson, I. H. (2001). Effects of attention and unilateral neglect on auditory stream segregation. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 115–127.

Carlyon, R. P., Plack, C. J., Fantini, D. A., & Cusack, R. (2003). Crossmodal and non-sensory influences on auditory streaming. *Perception, 32,* 1393–1402.

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.

Ciocca, V., & Darwin, C. J. (1993). Effects of onset asynchrony on pitch perception: Adaptation or grouping. *Journal of the Acoustical Society of America, 93,* 2870–2878.

Cooke, M., & Ellis, D. P. W. (2001). The auditory organization of speech and other sources in listeners and computational models. *Speech Communication, 35,* 141–177.

Culling, J. F., & Darwin, C. J. (1993). Perceptual separation of simultaneous vowels: Within and across-formant grouping by F0. *Journal of the Acoustical Society of America, 93,* 3454–3467.

Culling, J. F., & Darwin, C. J. (1994). Perceptual and computational separation of simultaneous vowels: Cues arising from low-frequency beating. *Journal of the Acoustical Society of America, 95,* 1559–1569.

Culling, J. F., & Summerfield, Q. (1995). Perceptual separation of concurrent speech sounds: Absence of across-frequency grouping by common interaural delay. *Journal of the Acoustical Society of America, 98,* 785–797.

Cusack, R., & Carlyon, R. P. (2004). Auditory perceptual organization inside and outside the laboratory. In J. G. Neuhoff (Ed.), *Ecological psychoacoustics* (pp. 16–48). Boston: Elsevier.

Cusack, R., Deeks, J., Aikman, G., & Carlyon, R. P. (2004). Effects of location, frequency region, and time course of selective attention on auditory scene analysis. *Journal of Experimental Psychology: Human Perception and Performance, 30,* 643–656.

Cusack, R., & Roberts, B. (2000). Effects of differences in timbre on sequential grouping. *Perception & Psychophysics, 62,* 1112–1120.

Dabelsteen, T., McGregor, P. K., Holland, J., Tobias, J. A., & Pedersen, S. B. (1997). The signal function of overlapping singing in male robins. *Animal Behaviour, 53,* 249–256.

Dabelsteen, T., McGregor, P. K., Shepherd, M., Whittaker, X., & Pedersen, S. B. (1996). Is the signal value of overlapping different from that of alternating during matched singing in Great Tits? *Journal of Avian Biology, 27,* 189–194.

Darwin, C. J. (1984). Perceiving vowels in the presence of another sound: Constraints on formant perception. *Journal of the Acoustical Society of America, 76,* 1636–1647.

Darwin, C. J. (1997). Auditory grouping. *Trends in Cognitive Sciences, 1,* 327–333.

Darwin, C. J. (2006). Contributions of binaural information to the separation of different sound sources. *International Journal of Audiology, 45,* S20–S24.

Darwin, C. J., & Carlyon, R. P. (1995). Auditory grouping. In B. C. J. Moore (Ed.), *Hearing* (pp. 387–424). New York: Academic Press.

Darwin, C. J., & Ciocca, V. (1992). Grouping in pitch perception: Effects of onset asynchrony and ear of presentation of a mistuned component. *Journal of the Acoustical Society of America, 91,* 3381–3390.

Darwin, C. J., Ciocca, V., & Sandell, G. J. (1994). Effects of frequency and amplitude modulation on the pitch of a complex tone with a mistuned harmonic. *Journal of the Acoustical Society of America, 95,* 2631–2636.

Darwin, C. J., & Hukin, R. W. (1998). Perceptual segregation of a harmonic from a vowel by interaural time difference in conjunction with mistuning and onset asynchrony. *Journal of the Acoustical Society of America, 103,* 1080–1084.

Darwin, C. J., & Sutherland, N. S. (1984). Grouping frequency components of vowels: When is a harmonic not a harmonic. *Quarterly Journal of Experimental Psychology Section A-Human Experimental Psychology, 36,* 193–208.

de Cheveigne, A., McAdams, S., Laroche, J., & Rosenberg, M. (1995). Identification of concurrent harmonic and inharmonic vowels: A test of the theory of harmonic cancellation and enhancement. *Journal of the Acoustical Society of America, 97,* 3736–3748.

Dent, M. L., Larsen, O. N., & Dooling, R. J. (1997). Free-field binaural unmasking in budgerigars (*Melopsittacus undulatus*). *Behavioral Neuroscience, 111,* 590–598.

Deutsch, D. (1974). An auditory illusion. *Nature, 251,* 307–309.

Dowling, W. J., & Fujitani, D. S. (1971). Contour, interval, and pitch recognition in memory for melodies. *Journal of the Acoustical Society of America, 49*(Suppl 2), 524–531.

Dowling, W. J., Lung, K. M., & Herrbold, S. (1987). Aiming attention in pitch and time in the perception of interleaved melodies. *Perception & Psychophysics, 41,* 642–656.

Ehret, G., & Gerhardt, H. C. (1980). Auditory masking and effects of noise on responses of the green treefrog (*Hyla cinerea*) to synthetic mating calls. *Journal of Comparative Physiology A, 141,* 13–18.

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. (2000). Spectral contrasts underlying auditory stream segregation in goldfish (*Carassius auratus*). *Journal of the Association for Research in Otolaryngology, 1,* 120–128.

Fay, R. R., & Popper, A. N. (2000). Evolution of hearing in vertebrates: The inner ears and processing. *Hearing Research, 149,* 1–10.

Feng, A. S., & Ratnam, R. (2000). Neural basis of hearing in real-world situations. *Annual Review of Psychology, 51,* 699–725.

Fernández-Juricic, E., Poston, R., De Collibus, K., Morgan, T., Bastain, B.,

Martin, C., et al. (2005). Microhabitat selection and singing behavior patterns of male house finches (*Carpodacus mexicanus*) in urban parks in a heavily urbanized landscape in the western U.S. *Urban Habitats, 3,* 49–69.

Festen, J. M. (1993). Contributions of comodulation masking release and temporal resolution to the speech-reception threshold masked by an interfering voice. *Journal of the Acoustical Society of America, 94,* 1295–1300.

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.

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.

Foote, A. D., Osborne, R. W., & Hoelzel, A. R. (2004). Whale-call response to masking boat noise. *Nature, 428,* 910.

Freyman, R. L., Balakrishnan, U., & Helfer, K. S. (2001). Spatial release from informational masking in speech recognition. *Journal of the Acoustical Society of America, 109,* 2112–2122.

Gaudrain, E., Grimault, N., Healy, E. W., & Bera, J.-C. (in press). Effect of spectral smearing on the perceptual segregation of vowel sequences. *Hearing Research.*

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, USA, 99,* 9021–9025.

Gerhardt, H. C. (1975). Sound pressure levels and radiation patterns of vocalizations of some North American frogs and toads. *Journal of Comparative Physiology, 102,* 1–12.

Gerhardt, H. C., & Bee, M. A. (2006). Recognition and localization of acoustic signals. In P. M. Narins, A. S. Feng, R. R. Fay, & A. N. Popper (Eds.), *Hearing and sound communication in amphibians* (Vol. 28, pp. 113–146). New York: Springer.

Gerhardt, H. C., & Huber, F. (2002). *Acoustic communication in insects and anurans: Common problems and diverse solutions.* Chicago: Chicago University Press.

Gerhardt, H. C., & Klump, G. M. (1988). Masking of acoustic signals by the chorus background noise in the green treefrog: A limitation on mate choice. *Animal Behaviour, 36,* 1247–1249.

Ghazanfar, A. A., & Logothetis, N. K. (2003). Neuroperception: Facial expressions linked to monkey calls. *Nature, 423,* 937–938.

Gockel, H., Carlyon, R. P., & Micheyl, C. (1999). Context dependence of fundamental-frequency discrimination: Lateralized temporal fringes. *Journal of the Acoustical Society of America, 106,* 3553–3563.

Gockel, H., Carlyon, R. P., & Plack, C. J. (2004). Across-frequency interference effects in fundamental frequency discrimination: Questioning evidence for two pitch mechanisms. *Journal of the Acoustical Society of America, 116,* 1092–1104.

Goense, J. B. (2004). *The effect of noise bandwidth and modulation on signal detection for single neurons in the frog auditory midbrain.* Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign.

Grafe, T. U. (2005). Anuran choruses as communication networks. In P. K. McGregor (Ed.), *Animal communication networks* (pp. 277–299). Cambridge, UK: Cambridge University Press.

Greenfield, M. D. (2005). Mechanisms and evolution of communal sexual displays in arthropods and anurans. *Advances in the Study of Behavior, 35,* 1–62.

Grimault, N. (2004). Sequential auditory scene analysis for hearing impaired subjects. *Revue de Neuropsychologie, 14,* 25–39.

Grimault, N., Bacon, S. P., & Micheyl, C. (2002). Auditory stream segregation on the basis of amplitude-modulation rate. *Journal of the Acoustical Society of America, 111,* 1340–1348.

Grimault, N., Micheyl, C., Carlyon, R. P., Arthaud, P., & Collet, L. (2000).

Influence of peripheral resolvability on the perceptual segregation of harmonic complex tones differing in fundamental frequency. *Journal of the Acoustical Society of America, 108,* 263–271.

Grimault, N., Micheyl, C., Carlyon, R. P., Arthaud, P., & Collet, L. (2001). Perceptual auditory stream segregation of sequences of complex sounds in subjects with normal and impaired hearing. *British Journal of Audiology, 35,* 173–182.

Grose, J. H., & Hall, J. W. (1992). Comodulation masking release for speech stimuli. *Journal of the Acoustical Society of America, 91,* 1042–1050.

Guilford, T., & Dawkins, M. S. (1991). Receiver psychology and the evolution of animal signals. *Animal Behaviour, 42,* 1–14.

Hall, J. W., Grose, J. H., & Mendoza, L. (1995). Across-channel processes in masking. In B. C. J. Moore (Ed.), *Hearing* (pp. 243–266). New York: Academic Press.

Hall, J. W., Haggard, M. P., & Fernandes, M. A. (1984). Detection in noise by spectro-temporal pattern-analysis. *Journal of the Acoustical Society of America, 76,* 50–56.

Hartmann, W. M., & Johnson, D. (1991). Stream segregation and peripheral channeling. *Music Perception, 9,* 155–184.

Hartmann, W. M., McAdams, S., & Smith, B. K. (1990). Hearing a mistuned harmonic in an otherwise periodic complex tone. *Journal of the Acoustical Society of America, 88,* 1712–1724.

Hawley, M. L., Litovsky, R. Y., & Culling, J. F. (2004). The benefit of binaural hearing in a cocktail party: Effect of location and type of interferer. *Journal of the Acoustical Society of America, 115,* 833–843.

Hill, N. I., & Darwin, C. J. (1996). Lateralization of a perturbed harmonic: Effects of onset asynchrony and mistuning. *Journal of the Acoustical Society of America, 100,* 2352–2364.

Hine, J. E., Martin, R. L., & Moore, D. R. (1994). Free-field binaural unmasking in ferrets. *Behavioral Neuroscience, 108,* 196–205.

Hofer, S. B., & Klump, G. M. (2003). Within- and across-channel processing in auditory masking: A physiological study in the songbird forebrain. *Journal of Neuroscience, 23,* 5732–5739.

Holt, M. M., & Schusterman, R. J. (2007). Spatial release from masking of aerial tones in pinnipeds. *Journal of the Acoustical Society of America, 121,* 1219–1225.

Hukin, R. W., & Darwin, C. J. (1995). Comparison of the effect of onset asynchrony on auditory grouping in pitch matching and vowel identification. *Perception & Psychophysics, 57,* 191–196.

Hulse, S. H. (2002). Auditory scene analysis in animal communication. *Advances in the Study of Behavior, 31,* 163–200.

Hulse, S. H., MacDougall-Shackleton, S. A., & Wisniewski, A. B. (1997). Auditory scene analysis by songbirds: Stream segregation of birdsong by European starlings (*Sturnus vulgaris*). *Journal of Comparative Psychology, 111,* 3–13.

Ison, J. R., & Agrawal, P. (1998). The effect of spatial separation of signal and noise on masking in the free field as a function of signal frequency and age in the mouse. *Journal of the Acoustical Society of America, 104,* 1689–1695.

Iverson, P. (1995). Auditory stream segregation by musical timbre: Effects of static and dynamic acoustic attributes. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 751–763.

Izumi, A. (2001). Auditory sequence discrimination in Japanese monkeys: Effect of frequency proximity on perceiving auditory stream. *Psychologia, 44,* 17–23.

Jensen, K. K. (2007). Comodulation detection differences in the hooded crow (*Corvus corone cornix*), with direct comparison to human subjects. *Journal of the Acoustical Society of America, 121,* 1783–1789.

Katti, M., & Warren, P. S. (2004). Tits, noise and urban bioacoustics. *Trends in Ecology & Evolution, 19,* 109–110.

Kidd, G. Jr., Mason, C. R., & Arbogast, T. L. (2002). Similarity, uncertainty, and masking in the identification of nonspeech auditory patterns. *Journal of the Acoustical Society of America, 111,* 1367–1376.

Kidd, G. Jr., Mason, C. R., & Dai, H. (1995). Discriminating coherence in spectro-temporal patterns. *Journal of the Acoustical Society of America, 97,* 3782–3790.

Kidd, G. Jr., Mason, C. R., Deliwala, P. S., Woods, W. S., & Colburn, H. S. (1994). Reducing informational masking by sound segregation. *Journal of the Acoustical Society of America, 95,* 3475–3480.

Klump, G. M. (1996). Bird communication in the noisy world. In D. E. Kroodsma & E. H. Miller (Eds.), *Ecology and evolution of acoustic communication in birds* (pp. 321–338). Ithaca, NY: Cornell University Press.

Klump, G. M. (2005). Evolutionary adaptations for auditory communication. In J. Blauert (Ed.), *Communication acoustics* (pp. 27–46). New York: Springer.

Klump, G. M., Kittel, M., & Wagner, E. (2001). Comodulation masking release in the Mongolian gerbil. *Abstracts of the Association for Research on Otolaryngology, 25,* 84.

Klump, G. M., & Langemann, U. (1995). Comodulation masking release in a songbird. *Hearing Research, 87,* 157–164.

Klump, G. M., & Larsen, O. N. (1992). Azimuthal sound localization in the European starling (*Sturnus vulgaris*): 1. Physical binaural cues. *Journal of Comparative Physiology A, 170,* 243–251.

Klump, G. M., & Nieder, A. (2001). Release from masking in fluctuating background noise in a songbird's auditory forebrain. *Neuroreport, 12,* 1825–1829.

Kroodsma, D. E., & Miller, E. H. (Eds.). (1996). *Ecology and evolution of acoustic communication in birds.* Ithaca, NY: Cornell University Press.

Kwon, B. J. (2002). Comodulation masking release in consonant recognition. *Journal of the Acoustical Society of America, 112,* 634–641.

Langemann, U., Gauger, B., & Klump, G. M. (1998). Auditory sensitivity in the great tit: Perception of signals in the presence and absence of noise. *Animal Behaviour, 56,* 763–769.

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 P. K. McGregor (Ed.), *Animal communication networks* (pp. 451–480). Cambridge, UK: Cambridge University Press.

AQ: 12 Langemann, U., & Klump, G. M. (). Detecting modulated signals in modulated noise: I. Behavioural auditory thresholds in a songbird.

Larsen, O. N., Dooling, R. J., & Michelsen, A. (2006). The role of pressure difference reception in the directional hearing of budgerigars (*Melopsittacus undulatus*). *Journal of Comparative Physiology A, 192,* 1063–1072.

Lewis, E. R., & Fay, R. R. (2004). Environmental variables and the fundamental nature of hearing. In G. A. Manley, A. N. Popper, & R. R. Fay (Eds.), *Evolution of the vertebrate auditory system* (Vol. 22, pp. 27–54). New York: Springer.

Lin, W. Y., & Feng, A. S. (2003). GABA is involved in spatial unmasking in the frog auditory midbrain. *Journal of Neuroscience, 23,* 8143–8151.

Lippmann, R. P. (1997). Speech recognition by machines and humans. *Speech Communication, 22,* 1–15.

Lohr, B., Wright, T. F., & Dooling, R. J. (2003). Detection and discrimination of natural calls in masking noise by birds: Estimating the active space of a signal. *Animal Behaviour, 65,* 763–777.

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.

Mason, A. C., Oshinsky, M. L., & Hoy, R. R. (2001). Hyperacute directional hearing in a microscale auditory system. *Nature, 410,* 686–690.

McFadden, D. (1987). Comodulation detection differences using noise-band signals. *Journal of the Acoustical Society of America, 81,* 1519–1527.

McGregor, P. K. (Ed.). (2005). *Animal communication networks.* Cambridge, UK: Cambridge University Press.

Micheyl, C., Bernstein, J. G., & Oxenham, A. J. (2006). Detection and F0 discrimination of harmonic complex tones in the presence of competing tones or noise. *Journal of the Acoustical Society of America, 120,* 1493–1505.

Micheyl, C., & Carlyon, R. P. (1998). Effects of temporal fringes on fundamental-frequency discrimination. *Journal of the Acoustical Society of America, 104,* 3006–3018.

Micheyl, C., Carlyon, R. P., Cusack, R., & Moore, B. C. J. (2005). Performance measures of auditory organization. In D. Pressnitzer, A. de Cheveigne, S. McAdams, & L. Collet (Eds.), *Auditory signal processing: Physiology, psychoacoustics, and models* (pp. 203–211). New York: Springer.

Micheyl, C., Carlyon, R. P., Gutschalk, A., Melcher, J. R., Oxenham, A. J., Rauschecker, J. P., et al. (2007). The role of auditory cortex in the formation of auditory streams. *Hearing Research, 229,* 116–131.

Micheyl, C., & Oxenham, A. J. (2007). Across-frequency pitch discrimination interference between complex tones containing resolved harmonics. *Journal of the Acoustical Society of America, 121,* 1621–1631.

Micheyl, C., Shamma, S., & Oxenham, A. J. (2007). Hearing out repeating elements in randomly varying multitone sequences: A case of streaming? In B. Kollmeier, G. M. Klump, V. Hohmann, U. Langemann, M. Mauermann, S. Uppenkamp, & J. L. Verhey (Eds.), *Hearing: From basic research to applications* (pp.). Berlin: Springer. **AQ: 13**

Micheyl, C., Tian, B., Carlyon, R. P., & Rauschecker, J. P. (2005). Perceptual organization of tone sequences in the auditory cortex of awake Macaques. *Neuron, 48,* 139–148.

Miles, R. N., & Hoy, R. R. (2006). The development of a biologically-inspired directional microphone for hearing aids. *Audiology and Neuro-Otology, 11,* 86–94.

Miller, G. A., & Heise, G. A. (1950). The trill threshold. *Journal of the Acoustical Society of America, 22,* 637–638.

Moore, B. C. J., Glasberg, B. R., & Peters, R. W. (1986). Thresholds for hearing mistuned partials as separate tones in harmonic complexes. *Journal of the Acoustical Society of America, 80,* 479–483.

Moore, B. C. J., & Gockel, H. (2002). Factors influencing sequential stream segregation. *Acta Acustica United with Acustica, 88,* 320–333. **AQ: 14**

Moore, B. C. J., Peters, R. W., & Stone, M. A. (1999). Benefits of linear amplification and multichannel compression for speech comprehension in backgrounds with spectral and temporal dips. *Journal of the Acoustical Society of America, 105,* 400–411.

Moore, B. C. J., Shailer, M. J., Hall, J. W., & Schooneveldt, G. P. (1993). Comodulation masking release in subjects with unilateral and bilateral hearing impairment. *Journal of the Acoustical Society of America, 93,* 435–451.

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.

Naguib, M. (1999). Effects of song overlapping and alternating on nocturnally singing nightingales. *Animal Behaviour, 58,* 1061–1067.

Naguib, M. (2005). Singing interactions in songbirds. In P. K. McGregor (Ed.), *Animal communication networks* (pp. 300–319). Cambridge, UK: Cambridge University Press.

Naguib, M., Fichtel, C., & Todt, D. (1999). Nightingales respond more strongly to vocal leaders of simulated dyadic interactions. *Proceedings of the Royal Society of London Series B-Biological Sciences, 266,* 537–542.

Naguib, M., & Todt, D. (1997). Effects of dyadic vocal interactions on other conspecific receivers in nightingales. *Animal Behaviour, 54,* 1535–1543.

Narins, P. M. (1982). Effects of masking noise on evoked calling in the Puerto Rican coqui (*Anura, Leptodactylidae*). *Journal of Comparative Physiology, 147,* 439–446.

Narins, P. M., & Zelick, R. (1988). The effects of noise on auditory processing and behavior in amphibians. In B. Fritzsch, M. J. Ryan, W. Wilczynski, T. E. Hetherington, & W. Walkowiak (Eds.), *The evolution of the amphibian auditory system* (pp. 511–536). New York: Wiley.

AQ: 15 Nelken, I., Jacobson, G., Ahdut, L., & Ulanovsky, N. (2000). Neural correlates of comodulation masking release in auditory cortex of cats. *European Journal of Neuroscience, 12,* 495–495.

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

Neuert, V., Verhey, J. L., & Winter, I. M. (2004). Responses of dorsal cochlear nucleus neurons to signals in the presence of modulated maskers. *Journal of Neuroscience, 24,* 5789–5797.

Nieder, A., & Klump, G. M. (2001). Signal detection in amplitude-modulated maskers. II. Processing in the songbird's auditory forebrain. *European Journal of Neuroscience, 13,* 1033–1044.

Otter, K., McGregor, P. K., Terry, A. M. R., Burford, F. R. L., Peake, T. M., & Dabelsteen, T. (1999). Do female great tits (*Parus major*) assess males by eavesdropping? A field study using interactive song playback. *Proceedings of the Royal Society of London Series B-Biological Sciences, 266,* 1305–1309.

Partan, S., & Marler, P. (1999). Behavior: Communication goes multimodal. *Science, 283,* 1272–1273.

Patricelli, G. L., & Blickley, J. L. (2006). Avian communication in urban noise: Causes and consequences of vocal adjustment. *Auk, 123,* 639–649.

Peake, T. M. (2005). Eavesdropping in communication networks. In P. K. McGregor (Ed.), *Animal communication networks* (pp. 13–37). Cambridge, UK: Cambridge University Press.

Peake, T. M., Terry, A. M. R., McGregor, P. K., & Dabelsteen, T. (2001). Male great tits eavesdrop on simulated male-to-male vocal interactions. *Proceedings of the Royal Society of London Series B-Biological Sciences, 268,* 1183–1187.

Peake, T. M., Terry, A. M. R., McGregor, P. K., & Dabelsteen, T. (2002). Do great tits assess rivals by combining direct experience with information gathered by eavesdropping? *Proceedings of the Royal Society of London Series B-Biological Sciences, 269,* 1925–1929.

Popper, A. N., & Fay, R. R. (1997). Evolution of the ear and hearing: Issues and questions. *Brain, Behavior and Evolution, 50,* 213–221.

Pressnitzer, D., & Hupe, J. M. (2006). Temporal dynamics of auditory and visual bistability reveal common principles of perceptual organization. *Current Biology, 16,* 1351–1357.

Pressnitzer, D., Micheyl, C., Sayles, M., & Winter, I. M. (2007). Responses to long-duration tone sequences in the cochlear nucleus. *Abstracts of the Thirtieth Annual Midwinter Research Meeting of the Association for* **AQ: 16** *Research in Otolaryngology 131.*

Rabin, L. A., & Greene, C. M. (2002). Changes to acoustic communication systems in human-altered environments. *Journal of Comparative Psychology, 116,* 137–141.

Ratnam, R., & Feng, A. S. (1998). Detection of auditory signals by frog inferior collicular neurons in the presence of spatially separated noise. *Journal of Neurophysiology, 80,* 2848–2859.

Richards, D. G., & Wiley, R. H. (1980). Reverberations and amplitude fluctuations in the propagation of sound in a forest: Implications for animal communication. *American Naturalist, 115,* 381–399.

Roberts, B., & Bailey, P. J. (1996a). Regularity of spectral pattern and its effects on the perceptual fusion of harmonics. *Perception & Psychophysics, 58,* 289–299.

Roberts, B., & Bailey, P. J. (1996b). Spectral regularity as a factor distinct from harmonic relations in auditory grouping. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 604–614.

Roberts, B., & Brunstrom, J. M. (1998). Perceptual segregation and pitch shifts of mistuned components in harmonic complexes and in regular inharmonic complexes. *Journal of the Acoustical Society of America, 104,* 2326–2338.

Roberts, B., & Brunstrom, J. M. (2001). Perceptual fusion and fragmentation of complex tones made inharmonic by applying different degrees of frequency shift and spectral stretch. *Journal of the Acoustical Society of America, 110,* 2479–2490.

Roberts, B., Glasberg, B. R., & Moore, B. C. J. (2002). Primitive stream segregation of tone sequences without differences in fundamental frequency or passband. *Journal of the Acoustical Society of America, 112,* 2074–2085.

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.

Rowe, C. (1999). Receiver psychology and the evolution of multicomponent signals. *Animal Behaviour, 58,* 921–931.

Rowe, C. (2002). Sound improves visual discrimination learning in avian predators. *Proceedings of the Royal Society of London Series B-Biological Sciences, 269,* 1353–1357.

Rowe, C., & Skelhorn, J. (2004). Avian psychology and communication. *Proceedings of the Royal Society of London Series B-Biological Sciences, 271,* 1435–1442.

Scheffers, M. (1983). *Sifting vowels: Auditory pitch analysis and sound segregation.* Groningen University, Groningen, The Netherlands. **AQ: 17**

Schooneveldt, G. P., & Moore, B. C. J. (1989). Comodulation masking release (CMR) as a function of masker bandwidth, modulator bandwidth, and signal duration. *Journal of the Acoustical Society of America, 85,* 273–281.

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.

Shinn-Cunningham, B. G., Ihlefeld, A., Satyavarta, & Larson, E. (2005). **AQ: 18** Bottom-up and top-down influences on spatial unmasking. *Acta Acustica United with Acustica, 91,* 967–979. **AQ: 19**

Shinn-Cunningham, B. G., Schickler, J., Kopco, N., & Litovsky, R. (2001). Spatial unmasking of nearby speech sources in a simulated anechoic environment. *Journal of the Acoustical Society of America, 110,* 1118–1129.

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.

Singh, N. C., & Theunissen, F. E. (2003). Modulation spectra of natural sounds and ethological theories of auditory processing. *Journal of the Acoustical Society of America, 114,* 3394–3411.

Slabbekoorn, H., & den Boer-Visser, A. (2006). Cities change the songs of birds. *Current Biology, 16,* 2326–2331.

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

Slabbekoorn, H., Yeh, P., & Hunt, K. (2007). Sound transmission and song divergence: A comparison of urban and forest acoustics. *Condor, 109,* 67–78.

Stickney, G. S., Zeng, F. G., Litovsky, R., & Assmann, P. (2004). Cochlear implant speech recognition with speech maskers. *Journal of the Acoustical Society of America, 116,* 1081–1091.

Summerfield, Q., & Assmann, P. F. (1991). Perception of concurrent vowels: Effects of harmonic misalignment and pitch-period asynchrony. *Journal of the Acoustical Society of America, 89,* 1364–1377.

Sun, J. W. C., & Narins, P. A. (2005). Anthropogenic sounds differentially affect amphibian call rate. *Biological Conservation, 121,* 419–427.

Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift für Tierpsychologie, 20,* 410–433.

Tsuzaki, M., Takeshima, C., Irino, T., & Patterson, R. D. (2007). Auditory stream segregation based on speaker size, and identification of size-modulated vowels. In B. Kollmeier, G. M. Klump, V. Hohmann, U. Langemann, M. Mauermann, S. Uppenkamp, & J. L. Verhey (Eds.), *Hearing: From basic research to applications* (pp.). Berlin: Springer.

Turnbull, S. D. (1994). Changes in masked thresholds of a harbor seal (*Phoca vitulina*) associated with angular separation of signal and noise sources. *Canadian Journal of Zoology, 72,* 1863–1866.

van Noorden, L. P. A. S. (1975). *Temporal coherence in the perception of tone sequences.* Doctoral dissertation, Eindhoven University of Technology, The Netherlands.

Van Valkenburg, D., & Kubovy, M. (2004). From Gibson's fire to Gestalts: A bridge-building theory of perceptual objecthood. In J. G. Neuhoff (Ed.), *Ecological psychoacoustics* (pp. 114–147). Boston: Elsevier.

Verhey, J. L., Pressnitzer, D., & Winter, I. M. (2003). The psychophysics and physiology of comodulation masking release. *Experimental Brain Research, 153,* 405–417.

Vliegen, J., Moore, B. C. J., & Oxenham, A. J. (1999). The role of spectral and periodicity cues in auditory stream segregation, measured using a temporal discrimination task. *Journal of the Acoustical Society of America, 106,* 938–945.

Vliegen, J., & Oxenham, A. J. (1999). Sequential stream segregation in the absence of spectral cues. *Journal of the Acoustical Society of America, 105,* 339–346.

Wang, D., & Brown, G. J. (Eds.). (2006). *Computational auditory scene analysis: Principles, algorithms, and applications.* Hoboken, NJ: Wiley-IEEE Press.

Warren, P. S., Katti, M., Ermann, M., & Brazel, A. (2006). Urban bioacoustics: It's not just noise. *Animal Behaviour, 71,* 491–502.

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.

Wells, K. D., & Schwartz, J. J. (2006). The behavioral ecology of anuran communication. In P. M. Narins, A. S. Feng, R. R. Fay, & A. N. Popper (Eds.), *Hearing and sound communication in amphibians* (Vol. 28, pp. 44–86). New York: Springer.

Wiley, R. H. (2006). Signal detection and animal communication. *Advances in the Study of Behavior, 36,* 217–247.

Wisniewski, A. B., & Hulse, S. H. (1997). Auditory scene analysis in European starlings (*Sturnus vulgaris*): Discrimination of song segments, their segregation from multiple and reversed conspecific songs, and evidence for conspecific song categorization. *Journal of Comparative Psychology, 111,* 337–350.

Wollerman, L. (1999). Acoustic interference limits call detection in a neotropical frog *Hyla ebraccata. Animal Behaviour, 57,* 529–536.

Wollerman, L., & Wiley, R. H. (2002). Background noise from a natural chorus alters female discrimination of male calls in a Neotropical frog. *Animal Behaviour, 63,* 15–22.

Wood, W. E., & Yezerinac, S. M. (2006). Song sparrow (*Melospiza melodia*) song varies with urban noise. *Auk, 123,* 650–659.

Yost, W. A. (1991). Auditory image perception and analysis: The basis for hearing. *Hearing Research, 56,* 8–18.

Yost, W. A. (1997). The cocktail party problem: Forty years later. In R. H. Gilkey & T. R. Anderson (Eds.), *Binaural and spatial hearing in real and virtual environments* (pp. 329–347). Mahwah, NJ: Erlbaum.

Zwicker, U. T. (1984). Auditory recognition of diotic and dichotic vowel pairs. *Speech Communication, 3,* 265–277.

AQ: 20

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES   1

AQ1: Author: Please add several key words.

AQ2: Author: Please provide reference citation for the McGregor article in this issue.

AQ3: Author: Is the change in order OK?

AQ4: Author: Has this been published or is it an in press cite?

AQ5: Author: Has then been published or is it an in press cite?

AQ6: Author: Please provide the date of the version that you read unless it has been published. In that case, provide the publication date.

AQ7: Author: Provide the date of the version that you read unless it has been published. In that case, provide the publication date.

AQ8: Author: Provide the date of the version that you read unless it has been published. In that case, provide the publication date.

AQ9: Author: Provide the date of the version that you read unless it has been published. In that case, provide the publication date.

AQ10: Author: Please provide the page number of the quote.

AQ11: Author: Provide the date of the version that you read. Has this been accepted for publication yet?

AQ12: Author: Provide the date of the version that you read. Has this been accepted for publication yet?

AQ13: Author: Has this been published? If not, the date should read "in press" not "2007"." If yes, please provide inclusive page numbers of the chapter.

AQ14: Editor: I have seen this journal cited as "Acta Acustica united with Acustica," "Acta Acustica United with Acustica," and "Acta Acustica-Acustica." Which does APA prefer?

AQ15: Author: Please check inclusive page numbers.

AQ16: Author: If these abstracts are published, please provide publisher name and location. Is 131 the page number?

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES   2

AQ17: Author: Is this a doctoral dissertation? Has it been published?

AQ18: Author: Please check the authors for this reference. The citation I found did not list Ihlefeld and had A. as the initial for Satyavarta.

AQ19: Editor: I have seen this journal cited as "Acta Acustica united with Acustica," "Acta Acustica United with Acustica," and "Acta Acustica-Acustica." Which does APA prefer?

AQ20: Author: Has this been published? If so, provide page numbers. If not, use "in press" not "2007" as the date.

AQ21: Author: Is this figure reprinted from Grafe?