



# HHS Public Access

Author manuscript

*Int J Psychophysiol.* Author manuscript; available in PMC 2016 February 01.

Published in final edited form as:

*Int J Psychophysiol.* 2015 February ; 95(2): 216–237. doi:10.1016/j.ijpsycho.2014.01.004.

## Treefrogs as Animal Models for Research on Auditory Scene Analysis and the Cocktail Party Problem

Mark A. Bee\*

Department of Ecology, Evolution and Behavior, University of Minnesota, 100 Ecology, 1987 Upper Buford Circle, St. Paul, MN 55108, U.S.A.

### Abstract

The perceptual analysis of acoustic scenes involves binding together sounds from the same source and separating them from other sounds in the environment. In large social groups, listeners experience increased difficulty performing these tasks due to high noise levels and interference from the concurrent signals of multiple individuals. While a substantial body of literature on these issues pertains to human hearing and speech communication, few studies have investigated how nonhuman animals may be evolutionarily adapted to solve biologically analogous communication problems. Here, I review recent and ongoing work aimed at testing hypotheses about perceptual mechanisms that enable treefrogs in the genus *Hyla* to communicate vocally in noisy, multi-source social environments. After briefly introducing the genus and the methods used to study hearing in frogs, I outline several functional constraints on communication posed by the acoustic environment of breeding “choruses”. Then, I review studies of sound source perception aimed at uncovering how treefrog listeners may be adapted to cope with these constraints. Specifically, this review covers research on the acoustic cues used in sequential and simultaneous auditory grouping, spatial release from masking, and dip listening. Throughout the paper, I attempt to illustrate how broad-scale, comparative studies of carefully considered animal models may ultimately reveal an evolutionary diversity of underlying mechanisms for solving cocktail-party-like problems in communication.

### Keywords

Auditory grouping; Auditory stream segregation; Comodulation masking release; Dip listening; Perceptual restoration; Spatial release from masking

---

© 2014 Elsevier B.V. All rights reserved.

\* Correspondence author at: University of Minnesota, Department of Ecology, Evolution and Behavior, 100 Ecology, 1987 Upper Buford Circle, St. Paul, MN 55108, U.S.A. Tel.: +1-612-624-6749; fax: +1-612-624-6777. mbee@umn.edu (M. Bee)..

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

## 1. Introduction

In scientific investigation, the smallest details of method are of the highest importance. The right choice of animal, an instrument constructed in a certain way, or the use of one reagent instead of another, often suffices to resolve the most important of general questions.

–Claude Bernard (1865)

For such a large number of problems there will be some animal of choice or a few such animals on which it can be most conveniently studied.

–August Krogh (1929)

...natural selection does not work as an engineer works. It works like a tinkerer – a tinkerer who does not know exactly what he is going to produce but uses whatever he finds around him...to produce some kind of workable object.

–François Jacob (1977)

### 1.1 The “right choice” of animal

At least three different principles guide the “right choice” of animal models for investigating problems of interest in modern biomedical and behavioral research. One is that the animal is evolutionarily closely related to humans. That is, we choose animals that are closely related to us because we believe them to be somehow more “like us.” Choices of animal models based on this principle assume, often tacitly, that organismal systems are evolutionarily conserved and respond to external and internal conditions, pharmaceutical agents, and social and environmental stressors similarly to how the same systems operate in humans. This assumption accounts, in part, for the widespread use of mammals in biomedical and behavioral research, though it does not always hold (Seok et al., 2013).

A second guiding principle relates to tractability and versatility. Animals with small body sizes, rapid development and short generation times, minimal housing needs, and for which powerful genetic tools can be applied are often favored animals for laboratory research. This principle, for example, has contributed to organisms like *Drosophila* and *C. elegans* becoming model systems of choice in diverse research fields. Other issues of tractability and versatility relate to more specific research requirements, such as the ease with which certain anatomical structures, tissues, or cells can be viewed, accessed, harvested, manipulated or cultured.

Finally, a third guiding principle, one sometimes known as Krogh's Principle (Krebs, 1975), is that animals are chosen as research models because they are ethological specialists adapted to solve particularly challenging problems shared in some way with humans (Miller, 2004). Consider two ethological specialists in sound source localization. Barn owls (*Tyto alba*) hunt in the dark of night by exploiting the rustling sounds of scurrying nocturnal prey. These owls exhibit remarkable sound localization abilities and have featured prominently in basic research in auditory neuroscience targeted at discovering neural mechanisms for sound localization (reviewed in Konishi, 2003). Females of the tiny parasitoid fly, *Ormia*

*ochracea*, are specialists at localizing singing male crickets in the dark. Their miniscule ears have become a model for producing highly directional hearing aid microphones (Miles and Hoy, 2006). The critical point to be made here is that barn owls and *Ormia* flies use *different mechanisms* to solve their respective localization problems. Owls employ neural processing involving delay lines and coincidence detectors, while *Ormia* relies on mechanically coupled ears. As these two examples illustrate, comparisons between distantly related taxa can provide critical opportunities to uncover evolutionary diversity in the mechanisms natural selection has contrived to solve common problems (Gerhardt and Huber, 2002; Schnupp and Carr, 2009). A second point concerning ethological specialists, like owls and flies, is that they are often more distantly related to humans than are many “traditional” mammalian models. Mongolian gerbils (*Meriones unguiculatus*) and domestic cats (*Felis catus*), for instance, are two mammals commonly used in studies of sound localization. They last shared a common ancestor with humans some 94.5 and 96.2 million years ago, respectively (Table 1). These divergence times are relatively recent considering that the last common ancestors humans shared with barn owls and *Ormia* lived some 322.4 and 725.5 million years ago, respectively (Hedges et al., 2006). Thus, focusing biomedical research on a small number of mammalian species because they are more closely related to (and hence “more like”) us potentially misses out on discovering how “evolution the tinkerer” has solved problems in other organisms. A comparative neuroethological approach is one that explicitly embraces the study of this potential diversity (Miller, 2004).

A problem of current research interest in auditory neuroscience is one that we humans share in common with a diversity of other animals: communicating acoustically in crowded and noisy social environments (Bee and Micheyl, 2008; Hulse, 2002). For humans, understanding speech becomes difficult in noisy group settings when multiple people are speaking simultaneously. This problem is aptly termed the “cocktail party problem” (Bronkhorst, 2000; Cherry, 1953; McDermott, 2009). The cocktail party problem represents a speech-specific example of the more general problem of auditory scene analysis (Bregman, 1990), whereby we perceptually bind together those sounds originating from a common source and separate them from the other sounds in the acoustic scene. Importantly, people with hearing impairments report increased difficulty understanding speech in noisy social settings, even when using hearing aids or cochlear implants (Moore, 2003). Likewise, computer algorithms for automated speech recognition perform relatively poorly in multi-talker situations (Cooke et al., 2010). Hence, understanding how healthy auditory systems have evolved to cope with these sorts of problems has important implications for both human health and computer technology.

Following Krogh's Principle, one important consideration guiding the “right choice” of animal model to study the cocktail party problem should be identifying potential ethological specialists that have evolved to solve biologically analogous problems. Fortunately, in a broad evolutionary sense, the cocktail party problem can be viewed as a speech-specific example of a much more general communication problem that is encountered, and solved, by many nonhuman animals (Bee and Micheyl, 2008; Hulse, 2002). Among the best-studied animal groups in this regard are dawn choruses and evening roosts of songbirds (Klump, 1996), large colonies of seabirds, such as penguins (Aubin and Jouventin, 1998, 2002), and

breeding choruses of frogs (Bee, 2012; Vélez et al., 2014) and insects (Römer, 2014). In the large ethological literature on animal communication, biological analogues of the human cocktail party problem have been more thoroughly investigated in frogs than in any other group of animals (Bee, 2012; Feng and Schul, 2007; Hulse, 2002; Narins and Zelick, 1988; Vélez et al., 2014).

## 1.2 Frogs as ethological specialists for communicating in noise

Many frogs form large and noisy social aggregations during their breeding seasons (Fig. 1; reviewed in Gerhardt and Huber, 2002; Wells, 2007). Quite commonly, these cacophonous choruses comprise many (perhaps a dozen or more) different species using the same habitat (e.g., a small pond or stream) at the same time. Within choruses, male frogs produce loud vocalizations termed “advertisement calls” to attract females for mating and to defend their calling sites and territories against rival males (reviewed in Gerhardt and Huber, 2002; Wells, 2007). Individual male frogs can produce calls reaching peak sound pressure levels of 95 to 105 dB pSPL (re 20  $\mu$ Pa) measured at a distance of 1 m (Gerhardt, 1975). (This is a rather remarkable feat considering that most frogs are quite small, with body sizes ranging from about 1.5 cm to 10 cm in length.) The sustained background noise levels in frog choruses can reach 75 to 85 dB SPL or more (Gerhardt and Klump, 1988; Narins, 1982; Schwartz et al., 2001; Swanson et al., 2007; M. S. Caldwell and M. A. Bee, unpublished data). In at least one species, choruses have been reported to be audible to humans from a distance of 2 km (Arak, 1983). Within the acoustic scene of a breeding chorus, a female frog must successfully accomplish a number of perceptual tasks in order to reproduce. She must detect a calling male, recognize him as a member of her own species based on analysis of the species-specific spectral and temporal properties of his calls, assess his potential quality as a mate, and localize him (Gerhardt and Bee, 2007). For their part, male frogs also recognize and estimate distances to competitive rivals of their own species, and in at least one species, males even learn to recognize the individual voice qualities of established territorial neighbors (Gerhardt and Bee, 2007). We know from decades of research that visual and chemical cues are not required for species or individual recognition and source localization in frogs (Gerhardt and Huber, 2002). Thus, frogs can be considered ethological specialists when it comes to solving cocktail-party-like problems in vocal communication.

There is a vast literature on hearing and sound communication in frogs (see reviews in Bee, 2012; Bee et al., 2013; Capranica, 1976; Fay and Simmons, 1999; Frishkopf and Goldstein Jr, 1963; Fritsch et al., 1988; Gerhardt and Huber, 2002; Kelley, 2004; Narins et al., 2007; Ryan, 2001; Schwartz and Bee, 2014; Vélez et al., 2014; Wilczynski and Capranica, 1984; Wilczynski and Ryan, 2010). This previous work provides the backdrop for this review. Here, I review recent and ongoing research with North American treefrogs in the genus *Hyla* aimed at understanding the mechanisms by which frogs perceive sound sources in noisy, multi-source environments. The review highlights research on three particularly well-studied species, the green treefrog (*Hyla cinerea*), Cope's gray treefrog (*Hyla chrysoscelis*), and the eastern gray treefrog (*Hyla versicolor*). Indeed, most of the work on hearing and sound communication in the genus *Hyla* has been conducted with these three species. In Section 2, I provide a brief account of why treefrogs represent an interesting animal model for studies of auditory scene analysis and the cocktail party problem from an evolutionary perspective.

Then in Section 3, I provide a brief overview of the most common methods used to study hearing and sound communication in frogs, with special emphasis on *Hyla*, as not all of these methods will be familiar to all readers. In this third section I also briefly consider the cocktail party problem from the treefrog's perspective, that is, in functional terms related to the communication biology of these animals. Then, in the main section of the review, Section 4, I describe research aimed at identifying the spectral, temporal, and spatial cues that frogs exploit to perceive discrete sound sources in multi-source environments. Finally, I end in Section 5 by looking forward to exciting opportunities for future research with these animals.

## 2. *Hyla* in an evolutionary context

Current estimates based on molecular and fossil evidence indicate frogs most recently shared a common ancestor with humans some 359.1 million years ago (Hedges et al., 2006). This divergence time is well after the evolutionary origin of hearing in vertebrates (Fay and Popper, 2000; Manley et al., 2004; Popper and Fay, 1997; Webster et al., 1992). This means that frogs and humans inherited their sense of hearing from their last common ancestor. However, there is compelling evidence that key features of the vertebrate auditory system, such as the tympanic middle ear, arose independently multiple times *after* the diversification of the major tetrapod lineages (Christensen-Dalsgaard and Carr, 2008). This means that the tympanic ear of frogs was derived independently of that in reptiles, birds, and mammals. While all frogs inherited their ears from a common amphibian ancestor, it is not entirely appropriate (though convenient as shorthand) to think of “frogs” as a single, homogeneous group. Consider the following perspective. Essentially modern frogs have been around for about 200 million years (Cannatella, 1995). The genus *Hyla* itself has been around some 40 to 50 million years (Wiens, 2007). Other frog models commonly used in auditory neuroscience research are in the genera *Rana*<sup>1</sup> (Hall, 1994) and *Xenopus* (e.g., Elliott et al., 2011). The lineages giving rise to modern *Hyla* and *Rana* diverged from their last common ancestor approximately 151.0 million years ago (Hedges et al., 2006). Both *Hyla* and *Rana* last shared a common ancestor with *Xenopus* about 198.6 million years ago (Hedges et al., 2006). Hence, more evolutionary time separates modern *Hyla*, *Rana*, and *Xenopus* from each other than separates humans from all other mammals used in modern hearing research (Table 1). Similar logic applies even within the genus *Hyla* itself. Cope's gray treefrogs (*H. chrysoscelis*) and eastern gray treefrogs (*H. versicolor*) are closely-related sister taxa that diverged from each other multiple times within the last 3–15 million years through a process involving repeated hybridization and genome duplication (Holloway et al., 2006; Ptacek et al., 1994; Smith et al., 2005, 2007). In addition, gray treefrogs and green treefrogs, *H. cinerea*, probably last shared a common ancestor about 25–33 million years ago (Smith et al., 2005, 2007). For comparison, a divergence in the range of 25–33 million years ago is similar to that estimated for the lineages leading to modern rhesus macaques and humans (Table 1).

---

<sup>1</sup>There has been recent controversy concerning frog taxonomy and nomenclature. Here and throughout, our use of nomenclature follows Pyron and Wiens (2011).

Obviously time itself is not a direct measure of evolutionary change. But longer time spans since divergence from a common ancestor provide greater opportunities for evolutionary mechanisms, such as mutation and natural selection, to produce novel traits that adapt organisms to their environments, including their acoustic environments. And this fact underlies the main take home point of this brief section: all organisms represent a mosaic of traits *shared* with their relatives due to inheritance from their most recent common ancestor, and *novel*, derived traits that make them different from other organisms. When it comes to frogs solving cocktail-party-like problems, there is thus potential for the operation of *both* shared mechanisms inherited from fish and common to all vertebrates (Fay and Popper, 2000; Popper and Fay, 1997) and novel mechanisms derived or elaborated in the lineages leading to different groups of modern frogs over the last 359 million years. This key insight from evolutionary biology has broad implications for why we should study auditory scene analysis in a diversity of animal models. It is the explanation for why nature can produce diverse solutions to common problems.

### 3. Brief primer on frog hearing

#### 3.1. Some methodological considerations

**3.1.1 Psychophysical methods**—In an ideal world it would be possible to train frogs to perform perceptual tasks using standard procedures that work well with many other animals, such as two-alternative forced choice and go/no-go procedures. Unfortunately, however, frogs have proven to be difficult subjects in traditional psychoacoustic experiments that require some form of conditioning. Alternative approaches, such as reflex modification, have been used successfully to study frog hearing (reviewed in Simmons and Moss, 1995). Though it has not yet been used widely with frogs, reflex modification has proven effective for deriving an audiogram and estimating critical ratios in green treefrogs, *H. cinerea* (Megela-Simmons et al., 1985; Moss and Simmons, 1986). By far, most behavioral studies of frog hearing and acoustic communication have relied on two natural behaviors, phonotaxis and evoked calling.

Female frogs in reproductive condition exhibit stereotyped patterns of walking or hopping toward sources of real or synthetic advertisement calls of their own species (Gerhardt, 1995; Rheinlaender and Klump, 1988). This approach toward sound is called *phonotaxis* (see Electronic Supplementary Material). The interpretation of a phonotaxis response is that the female has detected the signal, recognized it as the call of an appropriate mate, and localized it. When presented with multiple stimuli in choice tests, perceptual discrimination is indicated by differential phonotaxis toward one stimulus over another. Phonotaxis experiments represent powerful methods for investigating hearing in frogs because females reliably exhibit the behavior under highly controlled conditions in laboratory sound chambers. In many species, females can be repeatedly tested with little evidence for carry-over effects (e.g., Gerhardt et al., 2000b), making possible both large, within-subjects designs and the use of adaptive tracking procedures to estimate behavioral response thresholds (e.g., Bee and Schwartz, 2009). Phonotaxis has been used widely and for several decades to study the temporal and spectral selectivity of the auditory system in *Hyla* (reviewed in Gerhardt, 2001; Gerhardt and Huber, 2002). A second natural behaviour that



has also featured prominently in experimental studies of frog hearing is the *evoked vocal response* (Capranica, 1965). Male frogs commonly respond to hearing another nearby male, or one simulated by acoustic playback, by producing a response call of their own. Evoked calling has not been as widely used as phonotaxis to study aspects of hearing potentially related to sound source perception in the genus *Hyla* (but see Simmons et al., 1993).

**3.1.2. Physiological methods**—Much of the previous research investigating the physiological mechanisms of hearing in frogs has focused on relatively large-bodied species, such as those in the genus *Rana* (family Ranidae), including North American bullfrogs, *Rana catesbeiana*, northern leopard frogs, *R. pipiens*, and European common frogs, *R. temporaria* (reviewed in Fritzsche et al., 1988; Hall, 1994; Narins et al., 2007; Smotherman and Narins, 2000). Nevertheless, research on smaller hylids demonstrates that treefrogs too are tractable models for physiological studies of hearing. Several studies have used laser Doppler vibrometry to investigate the frequency response and directionality of the ears in *H. cinerea*, *H. chrysoscelis*, and *H. versicolor* (Caldwell et al., submitted; Jørgensen, 1991; Jørgensen and Gerhardt, 1991; Michelsen et al., 1986). Spontaneous otoacoustic emissions (SOAEs) have been recorded in all three of these species (van Dijk et al., 1996), and distortion product otoacoustic emissions (DPOAEs) have been reported in *H. cinerea* (van Dijk and Manley, 2001). The auditory brainstem response (ABR) has been recorded noninvasively in both *H. cinerea* (Buerkle et al., in review) and *H. chrysoscelis* (Schrode et al., in review) and invasive recordings of averaged evoked potentials have also been made in both species (Hillery and Fay, 1982; Mudry and Capranica, 1987). Extracellular single-unit and multi-unit recordings have been made from different stages of the ascending auditory pathway in these three *Hyla* species, including the auditory nerve (Ehret and Capranica, 1980; Ehret et al., 1983; Klump et al., 2004; Lim and Capranica, 1994), the superior olivary nucleus (Feng and Capranica, 1978), and the torus semicircularis, the likely homologue of mammalian inferior colliculus (Diekamp and Gerhardt, 1995; Hillery, 1984; Lombard and Straughan, 1974; Miranda and Wilczynski, 2009b; Penna et al., 1992). Whole-cell patch recordings have been made in the torus semicircularis in *H. chrysoscelis* and *H. versicolor* (Odom et al., 2012). Evoked potentials have also been recorded from the thalamus in *H. cinerea* (Mudry and Capranica, 1987). The influence of the endocrine system on auditory processing has been studied in *H. cinerea* (Miranda and Wilczynski, 2009a, b). The expression of immediate early genes by neurons throughout the auditory system has been used to investigate aspects of acoustic communication in some frogs (e.g., Burmeister et al., 2008; Hoke et al., 2004), though not yet in *Hyla*. Together, these studies illustrate the range of physiological tools currently available to investigate the mechanisms of auditory scene analysis in these animals.

### 3.2. Auditory processing of space, frequency, and time

In this section, I briefly discuss several features of auditory processing in frogs that are relevant to understanding the potential roles of various spatial, spectral, and temporal cues in auditory scene analysis. Readers are referred to the previous reviews cited above for more in-depth treatments of these topics. The auditory processes described in this section would generally be considered examples of low-level, bottom-up or “data-driven” processes that contribute to what Bregman (1990) has termed “primitive” auditory scene analysis. This

label is sometimes used to describe mechanisms that are automatic and obligatory. They are not dependent on higher-level processes related to attention, cognition, and learned expectations, processes that contribute to what Bregman (1990) termed “schema-based” scene analysis. While listener attention and expectations can have profound influences on auditory scene analysis in humans (Shamma et al., 2011; Snyder et al., 2012), the contribution of these higher-level cognitive processes to scene analysis by frogs has not been investigated (but see Farris and Ryan, 2011).

**3.2.1. Space**—Treefrogs are accurate in localizing sound sources in the azimuthal and vertical planes, as revealed in phonotaxis experiments with *H. cinerea* (Gerhardt and Rheinlaender, 1982; Klump et al., 2004; Rheinlaender et al., 1979), *H. chrysoscelis* (Caldwell and Bee, submitted), and *H. versicolor* (Jørgensen and Gerhardt, 1991), as well as the barking treefrog, *H. gratioiosa* (Klump and Gerhardt, 1989). Frogs generally lack external ear structures that could aid in sound localization, such as the pinnae of mammals and the facial ruffs and asymmetric ears of barn owls. Instead, the tympanic membranes sit flush with the sides of the head (see inset in Fig. 1A). Given their small size and the relatively low sound frequencies used for communication, the *external* binaural cues available for sound localization (i.e., inter-aural time and level differences) are generally thought to be too small to be useful. Like some insects and some other non-mammalian vertebrates, frogs have internally coupled ears that function, in part, as pressure difference receivers (reviewed in Christensen-Dalsgaard, 2005, 2011). Mechanical coupling of the air-filled middle-ear cavities through the mouth cavity and Eustachian tubes (Fig. 2A) functionally amplifies the magnitude of inter-aural level differences and imparts inherent directionality to each tympanum's response (Fig. 2B). This amplification and directionality result from the interaction of sound impinging on both the external and internal surfaces of each tympanum (Ho and Narins, 2006). However, the situation is made somewhat more complicated by the fact that sound need not reach the inner ear via tympanic pathways at all. Extratympanic pathways, for example through the body wall and lungs, can also transmit sound energy to the middle and inner ears of frogs (reviewed in Mason, 2007).

In *Hyla*, the directionality of the auditory system has been measured using laser Doppler vibrometry in *H. cinerea* (Michelsen et al., 1986), *H. chrysoscelis* (Fig. 2B; Caldwell et al., submitted), and *H. versicolor* (Jørgensen, 1991; Jørgensen and Gerhardt, 1991), recordings of single auditory nerve fibers in *H. cinerea* (Klump et al., 2004), and extracellular multi-unit recordings from the midbrain in *H. cinerea* (Rheinlaender et al., 1981) and *H. versicolor* (Schwartz and Gerhardt, 1995). The binaural response properties of single auditory neurons in the superior olivary nucleus have also been investigated in *H. cinerea* (Feng and Capranica, 1978).

**3.2.2. Frequency**—Amphibians are unique among vertebrates in having two different sensory papillae in each inner ear that encode airborne sound frequencies, the *amphibian papilla* (AP) and the *basilar papilla* (BP) (reviewed in Simmons et al., 2007). The AP is tonotopically organized and encodes low and intermediate frequencies (e.g., < 1.5 kHz), whereas the BP functions as a resonant structure that is broadly tuned to higher frequencies. The specific frequencies to which the AP and BP are tuned vary among species, but in all



species, one or both papillae are tuned to frequencies emphasized in the acoustic signals males produce, as revealed, for example, in single-unit recordings of auditory nerve fibers (Fig. 3; reviewed in Zakon and Wilczynski, 1988). Such observations led to the influential hypothesis that the frog auditory periphery functions as a “matched spectral filter” that selectively encodes frequencies in conspecific signals while filtering out other frequencies (Capranica and Moffat, 1983; Simmons, 2013)

Auditory sensitivity and frequency selectivity have been well studied in green treefrogs, *H. cinerea*. In this species, males produce advertisement calls with a bimodal frequency spectrum having energy near 0.64–1.34 kHz and 2.2–3.6 kHz (Figs. 1 & 3; Gerhardt, 2001). Females exhibit selective responses to advertisement calls with frequencies near 0.9 kHz and 2.7–3.5 kHz, preferring calls with simultaneous energy in both regions (reviewed in Gerhardt, 2001). Both behavioral (Megela-Simmons et al., 1985) and physiological (Buerkle et al., in review; Lombard and Straughan, 1974; Miranda and Wilczynski, 2009b; Penna et al., 1992) audiograms reveal increased sensitivity to these frequencies (Fig. 3A). Frequency selectivity is also highest in these regions, as demonstrated by lower critical ratios at frequencies corresponding to the spectral peaks of the call (Fig. 3A; Moss and Simmons, 1986). Auditory nerve fibers exhibit V-shaped tuning curves with values of  $Q_{10\text{dB}}$  (best frequency / bandwidth 10 dB above threshold) in the range of 1–4 (Fig. 3B; Capranica and Moffat, 1983; Ehret and Capranica, 1980; Ehret et al., 1983). As illustrated in Fig. 3B, many (but certainly not all) auditory nerve fibers are tuned to frequencies present in conspecific signals. Combination-sensitive neurons in the thalamus exhibit nonlinear facilitation in response to combinations of frequencies encoded by the AP and BP (Mudry and Capranica, 1987). By comparison with *H. cinerea*, spectral sensitivity and frequency selectivity have been less well studied in *H. chrysoscelis* and *H. versicolor*. What information is available for these two species, however, is entirely consistent with general patterns demonstrated for *H. cinerea* and other frogs (Bee, 2010; Gerhardt, 2005; Gerhardt and Huber, 2002; Gerhardt et al., 2007; Hillery, 1984; Schrode et al., 2012; Schrode et al., in review).

**3.2.3. Time**—Among the most behaviorally salient temporal properties of treefrog vocalizations are call duration, pulse rate or amplitude modulation (AM) rate, interpulse interval, pulse envelope shape, and pulse number (Gerhardt, 1988). Across species of *Hyla*, these temporal features variously encode information about species identity, aggressiveness, and genetic quality. In general, the frog periphery is not selective for these sorts of gross temporal features of sounds, instead reliably encoding a broad range of durations and AM rates. We also know that the anuran auditory nerve can encode the fine-temporal periodicity of harmonic sounds, such as the low (e.g., 100 Hz) fundamental frequency ( $F_0$ ) of the harmonically rich call of the bullfrog (*Rana catesbeiana*), even when the fundamental itself is missing (e.g., Schwartz and Simmons, 1990). Neural selectivity for temporal features of communication sounds emerges at higher levels of the auditory system in the form of a diversity of temporally tuned filters (reviewed in Hall, 1994). Many, but certainly not all, of these filters appear to be “matched temporal filters” for encoding the temporal features of conspecific calls (Rose and Capranica, 1984).

Behavioral studies of *H. chrysoscelis* and *H. versicolor* reveal a great deal of selectivity when it comes to processing the temporal features of their pulsatile advertisement calls,

especially pulse rate (Fig. 1; reviewed in Gerhardt, 2001; Gerhardt and Huber, 2002). However, the two species do not necessarily assign the same weight to the same acoustic features when evaluating pulse rate. Calls also differ between the two species in pulse duration, interpulse interval, and pulse rise time. Females of *H. chrysoscelis* strictly use only pulse rate as a temporal cue for species recognition. In contrast, although females of *H. versicolor* are behaviorally selective for conspecific pulse rates, they do not rely on rate, *per se*, but instead rely on assessments of absolute pulse duration and interpulse interval, as well as pulse rise time, in recognizing conspecific calls (Diekamp and Gerhardt, 1995; Gerhardt and Doherty, 1988; Gerhardt and Schul, 1999; Schul and Bush, 2002). Neural selectivity for these temporal acoustic properties in the midbrain appears to mirror that observed in behavioral studies (Diekamp and Gerhardt, 1995; Odom et al., 2012; Rose et al., 1985). These studies of the two gray treefrog species highlight the fact auditory processing can differ even among closely related species.

Among the more important recent findings related to temporal processing in frogs is the discovery of neurons in the midbrain that essentially “count” pulses (Alder and Rose, 1998; Edwards et al., 2002; Rose et al., 2011). These neurons integrate information over some threshold number of pulses separated by specific interpulse intervals. They fail to fire when an inadequate number of pulses is delivered or when the interval separating consecutive pulses is artificially increased. The properties of these neurons appear to differ between *H. chrysoscelis* and *H. versicolor* in ways that reflect species differences in behavioral selectivity for temporal call features (Odom et al., 2012).

### 3.3 Does “matched filtering” solve the frog's cocktail party problem?

Matched spectral and temporal filtering mechanisms are no doubt critical for filtering out the calls of other species in mixed-species choruses and for extracting important features of conspecific calls. To some, it might seem that matched filtering could provide all that is required for frogs to perceptually analyze acoustic scenes and hence avoid costly communication errors. But the matched filtering described in frogs is far from perfect. As illustrated in Fig. 3B, the frog's peripheral auditory system can encode a much wider range of frequencies than those present in conspecific vocalizations. Likewise, neurons in the central auditory system exhibit a broad range of temporal tuning properties, including low-pass, high-pass, band-pass, and band-stop neurons. It is not the case that even a majority of auditory neurons in the frog brain are selective for the temporal properties of conspecific vocalizations. Beyond these imperfections in the “match” of matched filtering, there is an additional and relatively straightforward explanation for why matched filtering cannot, by itself, solve the frog's cocktail party problem: in a dense chorus of conspecifics, *all* of the calls in a listener's acoustic scene would pass through the matched filters for conspecific calls. Hence, matched filtering may provide limited benefits in the context of perceptually segregating the individual calls of multiple, simultaneously signaling conspecifics in a chorus. So, how are individual conspecific males perceived as distinct sound sources? This question is taken up in Section 4 after briefly considering the biological consequences of communicating in a chorus.

### 3.4 Biological constraints imposed by noisy, multi-source environments

Rare is the human cocktail party where guests risk having sex with a different species or being eaten by a predator because they have trouble hearing. Not necessarily so for treefrogs. Laboratory studies of *H. cinerea*, *H. chrysoscelis* and *H. versicolor* have demonstrated the functional consequences of communicating in a chorus. High levels of background chorus noise and acoustic clutter resulting from call overlap can impair a female's ability to detect and recognize conspecific calls (Bee and Schwartz, 2009; Bee and Swanson, 2007; Ehret and Gerhardt, 1980; Gerhardt and Klump, 1988; Schwartz et al., 2013). Noise and acoustic clutter can also disrupt a female's ability to discriminate between different call types in the species' repertoire, such as advertisement calls versus aggressive calls (Schwartz and Gerhardt, 1989), between the calls of high-quality and low-quality mates (Bee, 2008b; Schwartz et al., 2001, 2008; Ward et al., 2013b), and even between conspecific and heterospecific calls (Bee, 2008a; Marshall et al., 2006; Schwartz, 1987; Schwartz and Gerhardt, 1995; Ward et al., 2013a). Hence, in choruses, females risk making communication errors ranging in severity from approaching an aggressive male to mating with a low-quality male or even a male of the wrong species. Evidence from field studies indicates such errors sometimes do, in fact, occur in nature. In mixed-species choruses of the two gray treefrogs, *H. chrysoscelis* and *H. versicolor*, for example, females occasionally mate with a male of the wrong species, a mistake that results in infertile or inviable offspring (Gerhardt et al., 1994). Chorus noise can also increase the time females require to choose a mate and decrease the directness of the path she takes to approach him, thereby increasing potential exposure to various predators in the natural environment (Caldwell and Bee, submitted). Together, the biological consequences of noise and acoustic clutter should act (or have acted in the past) as sources of evolutionary selection favoring auditory processing mechanisms that reduce these errors. Lest any reader should think frogs are somehow "bad" at solving cocktail-party-like problems, it is worth remembering that essentially modern frogs have been around for some 200 million years (Cannatella, 1995).

## 4. Auditory scene analysis in *Hyla*

In this section I review two main lines of research with *Hyla* related to how frogs perceive calling males as distinct sound sources in noisy, multi-source environments. The first line of research, on auditory grouping, has investigated the spectral, temporal, and spatial cues frogs use to perceptually bind signal components produced by the same male and to separate these sounds from other sounds in the acoustic scene. The second line of research has focused on the spatial and temporal cues the frog auditory system exploits to achieve a release from auditory masking by high levels of background noise. In studies of human sound source perception, a common metric of auditory grouping or the effects of masking release involves correct word or vowel recognition in various listening conditions (e.g., Bronkhorst, 2000; Darwin, 2008a). In frogs, biologically analogous perceptual tasks involve experimentally asking whether subjects recognize a particular acoustic signal as that of an appropriate mate or competitive rival, or whether they behaviorally discriminate between two or more signals. As illustrated in the studies reviewed in this section, synthetic models of frog calls and chorus noise have been extremely useful for probing the mechanisms of auditory grouping and masking release.

## 4.1 Auditory grouping

The analysis of acoustic scenes requires perceptually binding together the sounds produced by the same source that occur simultaneously across the frequency spectrum and sequentially through time. Grouping simultaneous sounds across frequency (e.g., harmonics, formants) has been termed “simultaneous integration,” whereas grouping temporally separated sounds through time (e.g., syllables, words, notes of a musical melody) is termed “sequential integration,” (Bregman, 1990). The cues for simultaneous and sequential auditory grouping have been well studied in humans using both speech and non-speech sounds (reviewed in Bregman, 1990; Carlyon, 2004; Darwin, 2005, 2008a; Darwin and Carlyon, 1995; Moore and Gockel, 2002, 2012). Due to the physics of sound production and propagation, sounds that are similar in frequency, harmonically related, start and stop at the same time, and originate from the same location in space are likely to have arisen from the same sound source. We humans exploit precisely these cues to construct percepts of the sound sources in acoustic scenes. As the experimental studies reviewed in this section indicate, it appears treefrogs exploit these same cues in auditory grouping. A common design theme in many of these experiments, most of which have been conducted with female subjects, is the use of stimuli that are only attractive (or are relatively more attractive) if their component parts are properly grouped together by the auditory system and segregated from other interfering sounds.

**4.1.1 Spectral proximity**—The role of spectral proximity in the perceptual organization of sound sequences has been well studied in humans using rapid, interleaved sequences of two repeating tones differing in frequency (e.g., “ABA...” or “ABAB...”) (e.g., Bregman, 1990; Carlyon, 2004; Moore and Gockel, 2002, 2012; van Noorden, 1975). When the spectral proximity between the A and B tones is high (i.e., small frequency separation,  $F$ ), listeners report hearing a single, integrated stream of tones bouncing up and down in frequency. If spectral proximity is reduced (i.e., when  $F$  becomes sufficiently large), listeners report hearing the interleaved sequence break apart into two separate “auditory streams” composed of either A or B tones alone (e.g., A–A–A– and –B–B–B). In general, when interleaved or overlapping sounds are sufficiently different in frequency, we tend to assign them to different auditory streams.

One study has investigated the role of spectral proximity in sequential integration in *Hyla* using an experimental paradigm inspired by the “ABAB...” design from previous human studies (Nityananda and Bee, 2011). Females of *H. chrysoscelis* were presented with a pulsatile “target” signal in the presence of a pulsatile “distractor” (Fig. 4A). The target repeated every few seconds to simulate a calling conspecific male, whereas the distractor played continuously. The pulses of both the target and distractor occurred with the species-specific pulse rate (~45.5 pulses/s). However, when the target was presented, its pulses were temporally interleaved with the pulses of the distractor. Thus, at times when the target signal came on, the composite pulse rate (target + distractor) was effectively doubled (~91 pulses/s) to a pulse rate that females strongly discriminate against (e.g., Schul and Bush, 2002). Across different trials, the  $F$  between the target and distractor was varied over a 1.5-octave range along a semitone musical scale. The prediction was that the target would only be attractive to subjects (and hence elicit phonotaxis) if its constituent pulses were grouped

together and perceptually segregated from the pulses in the distractor. Increasing  $F$  between 3 and 12 semitones resulted in a near linear increase in signal attractiveness (Fig. 4B). This result is thus consistent with the hypothesis that high spectral proximity (small  $F$ ) promoted perceptual integration (and hence an unattractive percept), whereas a decrease in spectral proximity (large  $F$ ) enabled perception of an integrated target that could be segregated from the distractor. The data also indicated that in mixed-species choruses, in which multiple species may call with pulsatile signals within the audible range, species differences in call frequency could promote segregation between the overlapping calls of conspecifics and heterospecifics. Whether magnitudes of  $F$  on the order of individual differences within a single species also promote stream segregation is an area of ongoing research.

**4.1.2 Harmonicity**—Many natural sounds, including speech and other animal vocalizations, possess rich harmonic structure that contributes to our perceptions of pitch and timbre. The relatedness of simultaneous spectral elements having frequencies that are integer multiples of a common fundamental frequency ( $F_0$ ) is termed “harmonicity,” and it represents a potent cue for assigning those sounds to the same auditory group in humans (reviewed in Bregman, 1990; Carlyon and Gockel, 2008; Darwin, 2005, 2008a; Darwin and Carlyon, 1995). Our reliance on harmonicity as a grouping cue is one of the reasons we hear complex tones comprised of multiple harmonics as a unified whole, instead of as a series of separate but simultaneous partials. The role of harmonicity as an auditory grouping cue has been widely studied in humans by “mistuning” specific harmonics to create varying degrees of inharmonicity. When mistuning is sufficiently large, the mistuned partial is no longer grouped with the remaining harmonics and can be heard as a separate but simultaneous sound.

Several behavioral studies have now investigated the influence of harmonicity on hearing and sound communication in the green treefrog, *H. cinerea*, and its closest relative, the barking treefrog, *H. gratiosa* (Bodnar, 1996; Gerhardt et al., 1990; Simmons, 1988; Simmons et al., 1993). Using reflex modification, Simmons (1988) elegantly showed that males of *H. cinerea* had masked thresholds for detecting a two-harmonic complex tone that were about 10 dB lower than their thresholds for detecting an inharmonic tone complex. These data convincingly demonstrated that the frog auditory system was sensitive to harmonicity. Follow-up studies of phonotaxis by females (Gerhardt et al., 1990) and evoked calling by males (Simmons et al., 1993), however, found that both sexes were insensitive to differences between harmonic and inharmonic calls comprised of two spectral components. These results contrast with those from a subsequent phonotaxis study of *H. gratiosa* (Bodnar, 1996). In this species, females were sensitive to differences between harmonic and inharmonic stimuli comprised of four spectral components, and preferences for harmonic versus inharmonic calls depended on the degree of frequency modulation in the signal. As discussed in more detail by Bodnar (1996), directly comparing the results from these studies is somewhat problematic given differences in the species investigated (*H. cinerea* versus *H. gratiosa*), differences in experimental approaches (e.g., reflex modification, phonotaxis, and evoked calling), and differences in the number of spectral components used to create stimuli (two versus four) and their relation to the tuning of the amphibian and basilar papillae.

Based on currently available data, it thus remains uncertain whether and to what extent harmonicity functions as an auditory grouping cue in *Hyla* (but see Simmons and Bean, 2000). Additional research on these questions is needed.

**4.1.3 Common onsets and offsets**—The human auditory system is adept at grouping simultaneous sounds in different frequency regions when they start and stop at the same time (reviewed in Bregman, 1990; Darwin, 2005, 2008a; Darwin and Carlyon, 1995). Our ability to exploit common onsets and offsets in auditory grouping plays an important role in vowel perception, for example. To my knowledge, there are currently no published studies that have examined common onsets and offsets as potential auditory grouping cues in frogs. Ongoing work with *H. cinerea* is testing the hypothesis that common onsets, common offsets, or both, function as cues that allow the frog auditory system to integrate spectrally remote components of signals into coherent auditory objects (Bee MA and Schrode KM, unpublished data). Recall that advertisement calls in this species have a bimodal spectrum with energy near 0.9 kHz and in the range of about 2.7–3.5 kHz (Fig. 1), that these two frequency regions are encoded by the amphibian and basilar papillae, respectively (Fig. 3), and that females prefer calls with simultaneous energy in both regions (reviewed in Gerhardt, 2001). In addition, combination sensitive neurons in the thalamus exhibit nonlinear facilitation when acoustic energy is present in both regions (Mudry and Capranica, 1987).

We are currently testing the hypothesis that the combination sensitivity demonstrated previously for both behavioral preferences and thalamic responses depends on time-critical integration of simultaneous onsets and offsets of sound energy in the bimodal spectrum (Bee MA and Schrode KM, unpublished data). In a preliminary study, we confirmed that females prefer calls having bimodal spectra with temporally coherent spectral peaks at 0.9 kHz and 2.7 kHz to unimodal alternatives with just one or the other of these two peaks by itself. In the critical manipulation, we are giving females a choice between (a) a synchronous bimodal call and (b) an asynchronous bimodal alternative in which we are simultaneously and systematically varying the asynchrony of the onsets and offsets of the two spectral peaks (Fig. 5A). As illustrated in Figure 5B, current trends in preliminary data suggest females use common onsets/offsets as an auditory grouping cue. At asynchronies smaller than 50 ms, for example, females treat synchronous and asynchronous stimuli the same (Fig. 5B). But at asynchronies of 50 ms and larger, the degree of preference for the synchronous signal over the asynchronous one approaches the same level of preference observed for calls with a bimodal spectrum over unimodal calls (Fig. 5B). In addition, ongoing experiments we are manipulating the asynchrony of onsets and offset independently as well as assessing discrimination between asynchronous bimodal calls and unimodal calls. While hardly conclusive, these preliminary data suggest the treefrog auditory system exploits common onsets, offsets, or both as cues for simultaneous integration. Additional work is needed to confirm this provisional finding and to investigate potential neural mechanisms underlying the observed patterns of behavior.

**4.1.4 Spatial proximity**—Spatial proximity can promote auditory grouping in humans, though it is perhaps not always as strong of a grouping cue as one might expect based on



personal introspection (reviewed in Darwin, 2008b). One study of *H. chrysoscelis* tested the hypothesis that spatial proximity functions as a cue for simultaneous integration (Bee, 2010). As in *H. cinerea*, the advertisement call in *H. chrysoscelis* has a bimodal spectrum (Fig. 1), and females prefer calls with bimodal spectra over unimodal alternatives (Bee, 2010; Gerhardt, 2005; Gerhardt and Doherty, 1988; Gerhardt et al., 2007). In a phonotaxis test, females were given a choice between (a) a spatially coherent bimodal call with peaks at 1.1 kHz and 2.2 kHz coming from the same speaker on one side of a circular test arena and (b) a spatially separated bimodal call in which the same two spectral peaks were broadcast from spatially separated speakers located on the other side of the arena (Fig. 6A). On different trials, the spatial proximity of the two speakers broadcasting the separated bimodal call were placed apart at angles ( $\theta$ ) of 7.5° (high proximity), 15°, 30° or 60° (low proximity). The prediction was that if spatial proximity promoted simultaneous integration, then females should increasingly prefer the spatially coherent bimodal call to the separated bimodal call as a function of increasing  $\theta$ . It was expected that preferences for the spatially coherent bimodal call would eventually increase to a level similar to the preference observed for bimodal calls over unimodal calls in control tests. The data did not exactly support this prediction, but for a surprising reason. Already at a separation of just 7.5°, preferences were shifted toward the spatially coherent bimodal call to levels not different from those expected based on discrimination between bimodal and unimodal controls (Fig. 6B). That is, even at 7.5° separation, females behaved as if they no longer grouped the two spectral peaks in the bimodal spectrum. Similar levels of preference for spatially coherent bimodal calls were maintained at angular separations of 15° to 60°. This somewhat surprising result suggests that, in fact, females of *H. chrysoscelis* may be quite sensitive to spatial coherence as a cue for simultaneous integration. It remains to be determined precisely which binaural disparities related to decreased spatial proximity (e.g., in level, arrival time, or ongoing phase) might have contributed most to the patterns of responses observed (see discussion in Bee, 2010).

The situation is different concerning the role of spatial coherence in sequential integration. Frogs appear to group temporally separated components of their signals over wide angles of spatial separation (e.g., Farris et al., 2002, 2005; Farris and Ryan, 2011; Gerhardt et al., 2000a). Thus far, two studies have tested the hypothesis that spatial proximity promotes sequential integration in gray treefrogs (*H. chrysoscelis* and *H. versicolor*) (Bee and Riemersma, 2008; Schwartz and Gerhardt, 1995). Both studies used interleaved sequences of pulses to take advantage of the natural selectivity that females have for conspecific pulse rates, which are about twice as fast in *H. chrysoscelis* (35–50 pulses/s) compared with *H. versicolor* (19–24 pulses/s) (Gerhardt, 2001). The basic experimental design is illustrated in Figure 7A. By interleaving two sequences of pulses, each with a pulse rate of, say, 20 pulses/s, it was possible to ask the animals whether they experienced an *integrated* percept corresponding to a single call with a pulse rate of 40 pulses/s (= one *H. chrysoscelis* call) or a *segregated* percept corresponding to two calls having pulse rates of 20 pulses/s (= two *H. versicolor* calls). In this way, the two studies were conceptually analogous to the use of “ABAB...” tone sequences in studies of auditory streaming in humans (e.g., Bregman, 1990; Moore and Gockel, 2002; van Noorden, 1975).

In their study of *H. versicolor*, Schwartz and Gerhardt (1995) gave females a choice between (a) two interleaved pulse sequences separated by just 5° on one side of the test arena and (b) two interleaved pulse sequences separated by either 45° or 120° on the opposite side of the arena (Fig. 7B). Because females of *H. versicolor* prefer calls with pulse rates near 20 pulses/s, the prediction was that reduced spatial proximity (45° or 120°) would promote segregation ( $\approx 2 \times 20$  pulses/s = *attractive*) and that high spatial proximity (5°) would promote integration ( $\approx 40$  pulses/s = *unattractive*). As the preference functions in Fig. 7C show, females preferentially chose calls separated by 120° over those separated by 5°, but there was no difference in the percentage of females choosing calls separated by 45° and 5°. These results indicated that a spatial separation larger than 45° was necessary for segregation, and that females similarly grouped pulse sequences separated by 5° and 45°. In follow-up tests of 5° versus 120°, Schwartz and Gerhardt (1995) systematically attenuated one of the two pulse sequences at 5° separation to estimate the influence of the auditory system's inherent directionality on the pattern of responses. Just 3 dB of attenuation was adequate to abolish the preference for calls separated by 120° (Fig. 7D).

Single-stimulus phonotaxis tests have been used to test the hypothesis that spatial proximity promotes perceptual integration in *H. chrysoscelis*, the gray treefrog that naturally prefers faster pulse rates (Bee and Riemersma, 2008). This time, however, the prediction was in the opposite direction. It was predicted that high spatial proximity would promote integration of two interleaved pulse sequences (25 pulses/s each) resulting in an *attractive* percept of a conspecific call (50 pulses/s). Spatial separation, then, was predicted to render calls relatively less attractive. On different trials, the two interleaved pulse sequences were separated by angles ( $\theta$ ) of 0°, 45°, 90°, and 180° (Fig. 7E). Response latencies increased (Fig. 7F) and the proportions of females responding decreased (Fig. 7G) as angular separation increased, as expected if spatially separated sequences were less attractive. However, as with *H. versicolor*, females of *H. chrysoscelis* appeared willing to group the two sequences over relatively large angles of separation. Large proportions of subjects (100%, 91%, and 74%) responded by eventually approaching one of the two spatially separated pulse sequences in the 45°, 90°, and 180° conditions, respectively (Fig. 7G).

Thus far, only one study has directly examined potential neural correlates underlying the role of spatial proximity as a grouping cue in *Hyla*. In their study of *H. versicolor*, Schwartz and Gerhardt (1995) used extracellular multi-unit recordings in the midbrain to investigate the extent to which the directionality of the auditory system might account for their behavioral results. Awake but immobilized animals were presented with two interleaved pulse sequences. One sequence was presented from a speaker on the ipsilateral side (relative to the recording site) and the other from a speaker on the contralateral side. Across stimulus presentations, the two speakers were separated by either 45° or 120° centered around the midline. Somewhat surprisingly, the difference in normalized response strength to interleaved sequences arising from the contralateral and ipsilateral sides of the animal was 2.1 dB at the 45° separation and 9.6 dB at the 120° separation. The magnitude of 9.6 dB observed at 120° was considerably larger than the 3 dB estimated for this angle in behavioral tests. Schwartz and Gerhardt (1995) discuss several possible hypotheses for this apparent discrepancy between behavior and physiology, most of which remain to be tested. One

additional hypothesis is that pulse interval counting neurons in the frog midbrain, reported first by Alder and Rose (1998) and later confirmed to also exist in gray treefrogs (Odom et al., 2012), have large spatial receptive fields that render the animals relatively unselective for spatial proximity in sequential integration tasks.

**4.1.5 Perceptual restoration**—In humans, an important part of perceiving sound sources turns out to be our ability to “fill in” missing information masked by intermittent loud noises (e.g., a door slamming or someone coughing). For example, we can perceive speech and other non-speech signals (e.g., frequency modulated glides) as if they continue through brief (e.g. 50–200ms) interruptions by loud noises, even if a silent gap is introduced into the signal at the time when the noise occurs (reviewed in Bregman, 1990; Kashino, 2006; King, 2007; Warren, 1999). Noise in the gaps evokes a so-called “continuity illusion” resulting from a process more generally known as “perceptual restoration” (also “temporal induction” or “auditory induction”). It is as if the noise induces illusory percepts, causing us to hear things that simply are not present in the acoustic stimulus. Two studies of gray treefrogs have now tested the hypothesis that frogs too experience perceptual restoration (Schwartz et al., 2010; Seeba et al., 2010). Both studies tested the specific hypothesis that females experience illusory pulses when actual pulses were removed from advertisement calls and replaced by short bursts of noise.

In two-stimulus choice tests with *H. chrysoscelis*, females were given a choice between (a) an attractive “standard call” with 35 pulses and (b) a “gap call” in which every other group of five pulses was replaced by a silent gap of equivalent duration (Fig. 8A; Seeba et al., 2010). The silent gaps introduced in the signal rendered it completely unattractive relative to the continuous standard call, which was unanimously preferred (Fig. 8A). Next, females were given a choice between (a) a gap call and (b) the same gap call with noise filling each silent gap (Fig. 8B). In this test, females unanimously preferred the “gap-filled call.” Naïve human listeners associated with conducting this particular experiment informally reported hearing illusory pulses in the gap-filled call (Bee, M. A. unpublished data; Electronic Supplementary Material). However, the question remained as to whether the frogs themselves actually experienced illusory pulses during the gaps filled with noise. To address this key question, the experimenters took advantage of the fact that females have strong preferences for longer calls having more pulses (Bee, 2008b; Gerhardt et al., 1996), a result they also confirmed in their study (Fig. 8C). In the critical test, females were given a choice between (a) a short, 20-pulse call that was complete and continuous and (b) a gap-filled call that also had 20 pulses, but had the total duration of a 35-pulse call because of the noise-filled gaps (Fig. 8D). If females actually heard illusory pulses during the noise, the prediction was that they should prefer the longer stimulus to the shorter stimulus. But this was not the case. In fact, there was a non-significant trend for females to prefer the shorter 20-pulse call nearly 2:1. Schwartz et al. (2010) have conducted analogous tests with *H. versicolor* using very similar methods to those used by Seeba et al. and found strikingly parallel results. Hence, these data from the two gray treefrogs did not support the hypothesis that the noise filling the gaps created the same illusory sensation of hearing the missing pulses for the frogs as it did for human listeners. It would be interesting to compare how the

pulse interval counting neurons in the gray treefrog midbrain respond to continuous calls, gap calls, and gap-filled calls.

**4.1.6 Summary of auditory grouping**—In Section 4.1, I have reviewed studies of the cues that allow treefrogs to construct meaningful auditory groups that correspond to communication signals. The evidence indicates treefrogs and humans rely on many of the same cues for auditory grouping. In terms of sequential integration, sounds (e.g., pulses) with high spectral proximity are grouped together, and those more separated in frequency are perceptually segregated. The treefrog auditory system is at least sensitive to harmonic versus inharmonic structure, suggesting this acoustic cue could function in simultaneous auditory grouping. Preliminary evidence suggests common onsets, offsets, or both likely function to promote simultaneous auditory grouping. The spatial proximity of sounds originating from the same location appear to strongly influence simultaneous auditory grouping, whereas treefrogs are fairly permissive of spatial separation in grouping sounds through time. Thus far, it appears that treefrogs do not integrate information from signals with intermittent loud noise to perceptually reconstruct missing or masked signal elements. In the next section, I turn to research on the question of how the treefrog auditory system segregates communication signals from the high levels of background noise present in chorus environments.

## 4.2 Masking release

Human speech intelligibility, typically measured in word recognition tasks, can improve markedly when sources of target speech and noise (e.g., broadband noise, speech-shaped noise, multi-talker babble, or competing sources of speech) are spatially separated in azimuth (Arbogast et al., 2002; Best et al., 2013; Bronkhorst and Plomp, 1988, 1992; Dirks and Wilson, 1969; Freyman et al., 1999; Hawley et al., 1999; Hirsh, 1950; Peissig and Kollmeier, 1997; Plomp and Mimpen, 1981; Shinn-Cunningham et al., 2001, 2005). This so-called “spatial release from masking” helps us to communicate in noisy social situations (Bronkhorst, 2000). Temporal features of noise can also influence the degree of auditory masking of speech. The level of background noise generated by multiple, simultaneous talkers is not constant, but rather fluctuates in level through time. The rate of these temporal fluctuations is slow (e.g., 3–6 Hz), reflecting the temporal structure of human speech (Houtgast and Steeneken, 1985). Human listeners are able to exploit these level fluctuations to catch brief “acoustic glimpses” of target speech at times when the background noise momentarily “dips” to a low level (Bacon et al., 1998; Festen and Plomp, 1990; Füllgrabe et al., 2006; Gustafsson and Arlinger, 1994). As a result, speech intelligibility is generally improved in temporally fluctuating maskers compared with noise lacking temporal envelope modulations. Such improvements are attributed to a phenomenon known as “dip listening.” Under some conditions, however, the temporal modulations present in noise can also impair auditory perception through phenomena known as “modulation masking” (Bacon and Grantham, 1989) or “modulation interference” (Kwon and Turner, 2001). There is some evidence to suggest that both dip listening and modulation interference can even operate simultaneously in humans (Kwon and Turner, 2001). In the next two sections, I review studies that have used phonotaxis experiments to investigate spatial release from masking and dip listening in *Hyla*.

**4.2.1 Spatial release from masking**—Schwartz and Gerhardt (1989) were the first to investigate spatial release from masking in frogs. In their two-stimulus choice tests with *H. cinerea*, females were given a choice between (a) an advertisement call on one side of the test arena and (b) an aggressive call on the other. In quiet, females discriminate between the two types of calls, favoring advertisement calls over aggressive calls, which possess additional temporal modulation. When both types of calls were presented near sources of broadband noise, behavioral discrimination between the two call types was impaired. Subjects experienced about 3 dB of masking release when signals and noises were spatially separated by 90° compared with a co-localized condition. However, spatial unmasking did not restore subjects' behavioral discrimination between the two call types. Subsequent studies of spatial release from masking have focused on phonotaxis in females of *H. chrysoscelis* (Bee, 2007, 2008a; Nityananda and Bee, 2012; Ward et al., 2013a). An initial study (Bee, 2007) based on reaction times measured how long subjects required to reach the sound source in the presence of “chorus-shaped noise” (i.e., noise shaped to have a long-term spectrum simulating natural choruses). Normalized reaction times became faster as the SNR increased between -12 dB and +12 dB. Importantly, there was a significant advantage when subjects were tested in a 90° spatially separated condition compared with a co-localized condition (Fig. 9A). The magnitude of this advantage was estimated to be at least 6 dB based on the separation between reaction-time functions (Fig. 9A). In a subsequent experiment (Bee, 2008a), spatial release from masking also provided an advantage when females discriminated between conspecific and heterospecific (*H. versicolor*) calls in two-stimulus choice tests (Fig. 9B). This improvement came about because spatial unmasking improves discrimination along a continuum of pulse rates ranging between average values for conspecific and heterospecific calls (Fig. 9C; Ward et al., 2013a).

Given the inherent directionality of their pressure-difference ears, the existence of extra-tympanic sound transmission pathways, and the uniqueness of two sensory papillae in each inner ear, the mechanisms by which frogs achieve spatial release from masking could differ in important ways from the mechanisms used by mammals. We are only now beginning to tease apart these issues. A recent study of *H. chrysoscelis* investigated whether the magnitude of spatial release from masking depends on the frequency content of the signal (Nityananda and Bee, 2012). “Signal-recognition thresholds,” which are conceptually analogous to “speech-reception thresholds” in human studies (Bee and Schwartz, 2009), were measured in response to advertisement calls presented in the presence of co-localized or separated (90°) chorus-shaped noise. The key manipulation was the use of three different synthetic signals. One signal had the normal bimodal spectrum of advertisement calls, and the two other signals had either just the lower or the higher spectral peak (Fig. 9D). The lower and higher spectral peaks are primarily encoded by the amphibian papilla and basilar papilla, respectively (Gerhardt, 2005; Hillery, 1984; Schrode et al., in review). Thus, the study was designed to test whether spatial release from masking varied depending on whether signals were primarily encoded by the amphibian papilla or the basilar papilla. Somewhat surprisingly, there was no evidence that the magnitude of spatial release varied significantly with the frequency spectrum of the signal (Fig. 9D). Masking release was on the order of 3 dB averaged across all three signal conditions. The magnitude of masking release was only slightly smaller in responses to the higher-peak signal (~2 dB) compared

with the bimodal signal (~3 dB) and the lower-peak signal (~4 dB). These results are important because they indicate that spectral integration of the two peaks in advertisement calls is unnecessary to achieve a benefit from spatial separation between sources of signals and noise.

Ongoing work with *H. chrysoscelis* is using laser Doppler vibrometry to measure the magnitude of spatial release from masking directly at the level of tympanum vibrations. Physical manipulations of the pressure-difference mechanism (e.g., damping contralateral tympanic input or blocking contralateral Eustachian tubes) should help identify precisely how the inherent directionality of the frog's pressure-difference ears contributes to spatial release from masking. Ultimately, the goal of this work is to better understand the sensory inputs to the periphery and the neural mechanisms that transform this input into a perceptual advantage in separated listening conditions. Related electrophysiological studies of the much larger northern leopard frogs, *R. pipiens*, have found average magnitudes of spatial unmasking on the order of 3 dB in auditory nerve fibers that are increased to 9 dB in the midbrain, in part due to GABAergic inhibition shaping the spatial selectivity of midbrain neurons (Lin and Feng, 2001, 2003; Ratnam and Feng, 1998).

Two issues regarding previous behavioral studies of spatial release from masking in frogs are worth briefly noting here. First, the reported magnitudes of spatial release from masking on the order of 3–6 dB are somewhat smaller than has been reported in some studies of humans (e.g., 6–10 dB; reviewed in Bronkhorst, 2000; but see e.g., Best et al., 2013). An important question, then is whether this magnitude of masking release is biologically relevant. The answer appears to be yes. As illustrated in Fig. 10, phonotaxis studies of level discrimination indicate differences as small as 2–4 dB, in both the absence and presence of noise, can shift female preferences in *H. chrysoscelis* (Bee et al., 2012). Similar results have been reported for *H. cinerea* (Gerhardt, 1987) and *H. versicolor* (Fellers, 1979). A second issue pertaining to all of the aforementioned behavioral studies of spatial release from masking in frogs is that they have involved phonotaxis experiments in which females are required to move toward a sound source. Hence, the SNR ratio is not constant throughout a trial (see Ward et al., 2013a). The implications of this important methodological limitation on estimating magnitudes of spatial release from masking have yet to be addressed. Alternative behavioral methods, such as reflex modification (Simmons and Moss, 1995), might prove useful in eliminating such issues in future studies.

**4.2.2 Dip listening**—Temporal fluctuations in level are a general feature of natural sounds and soundscapes, including the noise generated in frog choruses (Nelken et al., 1999; Richards and Wiley, 1980; Singh and Theunissen, 2003; Vélez and Bee, 2010, 2011). Slow rates of fluctuation (e.g., 5 Hz) dominate the temporal-modulation spectra of recordings of treefrog choruses (Fig. 11; Vélez and Bee, 2010, 2011). These slow modulations correspond to species-specific calling rates and call timing interactions among males in the chorus. Secondary peaks in these modulation spectra are also present at higher rates that correspond to the pulse rates of calls in choruses dominated by species with pulsatile calls (Fig. 11B). Thus, just like human speech has a characteristic temporal-modulation spectrum reflecting the temporal structure of speech (Houtgast and Steeneken, 1985), the background “noise” in



frog choruses has a temporal-modulation signature that reflects the calling behavior and temporal call structure of the species present in the chorus.

In a study comparing signal-recognition thresholds in the presence of unmodulated and sinusoidally amplitude modulated (SAM) chorus-shaped noises, females of *H. chrysoscelis* experienced about 2–4 dB of masking release in slowly fluctuating (< 5 Hz) SAM maskers (Fig. 11C; Vélez and Bee, 2011). These results were in agreement with those of a control study in which a threshold of 6 to 9 pulses was required to elicit positive phonotaxis in quiet conditions. Only the dips present in slowly (< 5 Hz) fluctuating SAM maskers were sufficiently long (measured at their 6-dB down points) to allow a threshold number of pulses to “fit” in a dip. Subsequent work has revealed a similar magnitude of masking release in noise with a slowly and randomly fluctuating envelope (Vélez and Bee, 2013). Together, these results suggest that acoustic glimpses of pulses afforded by dip listening can occur in the presence of slowly fluctuating noise backgrounds. In the study using SAM maskers, rates of fluctuation between 5 Hz and 20 Hz had no effect on signal-recognition thresholds (Fig. 11C). However, when the SAM masker fluctuated at faster rates more similar to the pulse rate of male advertisement calls (40–50 pulses/s), subjects experienced about 4–6 dB of additional masking (Fig. 11C; Vélez and Bee, 2010, 2011). This result suggests modulations present in the masker interfered with perception of the temporal pulse structure of the call, which is a critical acoustic cue for species recognition (Schul and Bush, 2002).

The observation that thresholds improved in slowly fluctuating noise, but deteriorated in rapidly fluctuating noise, raised the following question. Which of two processes, dip listening or modulation masking, might dominate in natural acoustic scenes? A detailed acoustical analysis of recordings of *H. chrysoscelis* choruses indicated that “dips” in level were relatively free from the faster modulations corresponding to the pulse rates of individual calls. This finding suggested that dip listening might function free from the influences of modulation masking when the background noise in natural chorus environments dipped to low levels (Vélez and Bee, 2011). However, this prediction was not supported in subsequent tests in which the maskers were “chorus-amplitude modulated” (CAM) noises created to have both the long-term frequency spectrum and temporal-modulation spectrum of natural chorus sounds (Vélez and Bee, 2013). This result would appear to contrast with that recently reported by Schwartz et al. (2013) for *H. versicolor*. In that study, subjects had lower masked signal-recognition thresholds in CAM noise compared with a steady state noise. However, this apparent discrepancy may be explained by an important methodological difference between the studies by Vélez and Bee (2013) and Schwartz et al. (2013). In the former study, modulated and unmodulated noises were equalized to have the same long-term equivalent RMS amplitudes ( $LC_{eq}$ ), whereas in the latter study, noises were equalized to have the same peak amplitudes. Hence, in the study by Schwartz et al. (2013), the CAM masker had a lower overall RMS amplitude compared with the steady-state masker, a fact that potentially explains the apparent masking release in modulated noise in that study.

While females of *H. chrysoscelis* benefited from dip listening in the presence of slowly fluctuating SAM maskers (Vélez and Bee, 2011), this was not the case in a study of *H. cinerea* (Vélez et al., 2012). Compared with signal-recognition thresholds measured in the

presence of steady-state, chorus-shaped noise, thresholds ranged from about 1.5 dB lower to 2.5 dB higher in the presence of SAM maskers modulated at rates between 0.625 Hz and 80 Hz (Fig. 11C). There was no indication that dip listening was occurring in this species, a result since corroborated in a subsequent study using both randomly fluctuating noise envelopes and CAM noises (Vélez and Bee, 2013). These contrasting results between *H. chrysoxcelis* and *H. cinerea* highlight the potential for species differences in effects related to dip listening and modulation masking. At present, we do not have any satisfactory explanation for this species difference.

The most recent work on dip listening in *H. chrysoxcelis* has investigated pulse-number discrimination in modulated and unmodulated chorus-shaped noise. Females of *H. chrysoxcelis* prefer males that produce relatively longer calls having more pulses, but chorus noise can interfere with the expression of these preferences (Bee, 2008b; Gerhardt et al., 1996; Schwartz et al., 2001; Ward et al., 2013b). In the sister-species, *H. versicolor*, such preferences are evolutionarily adaptive because the males that are able to produce longer calls also sire better-quality offspring (Welch et al., 1998). Vélez et al. (2013) recently tested the hypothesis that pulse-number discrimination might be better in modulated noise compared with unmodulated noise. The modulated noises consisted of SAM maskers that fluctuated at slow (0.625 Hz) or fast (45 Hz) rates and CAM maskers, in which the envelope fluctuated with the modulations typical of natural choruses (Fig. 11B). While preferences for longer calls were reduced in noise, there was no difference in pulse-number discrimination between unmodulated and modulated noise. Hence, these results failed to provide support for the hypothesis that dip listening improves discrimination between signals differing along an important biological continuum.

Ongoing work with *H. chrysoxcelis* is investigating the extent to which the frog auditory system might exploit the spectrotemporal correlation structure of natural chorus noise. Natural sounds, including frog choruses, exhibit correlated temporal fluctuations (“comodulation”) across the frequency spectrum (Klump, 1996; Nelken et al., 1999). Studies of humans (reviewed in Verhey et al., 2003) and a few other animals (Branstetter and Finneran, 2008; Branstetter et al., 2013; Fay, 2011; Jensen, 2007; Klump and Langemann, 1995; Langemann and Klump, 2007) indicate that auditory systems are sensitive to these correlations. Typically lower signal detection thresholds are found in comodulated noise. Preliminary studies of *H. chrysoxcelis* indicate that females are sensitive to the spectrotemporal correlation structure of masking noises, with relatively lower signal-recognition thresholds in comodulated noise conditions (Bee and Vélez, 2008; Lee et al., 2013). At a neurophysiological level, correlates of improved signal detection in comodulated noise have been reported for northern leopard frogs, *R. pipiens* (Goense and Feng, 2012). Whether improved performance in comodulated noise represents the contribution of mechanisms that operate within single auditory filters or across channels (e.g., across physically distinct sensory papillae in the inner ear) is a topic of current investigation.

**4.2.3 Summary on masking release**—Like humans, treefrogs benefit in signal recognition and discrimination tasks when sources of signals and noise are spatially separated. Some treefrogs, like humans, also benefit from listening in the dips of slowly fluctuating noise. Moreover, the magnitudes of masking release observed in behavioral

studies of treefrogs (e.g., 2–6 dB) and physiological studies of other frog species (e.g., 3–9 dB) are of potential biological relevance to the animals themselves. More work is needed to determine whether and to what extent mechanisms that enable dip listening in some species in the presence of some fluctuating noise backgrounds contribute to masking release in the natural listening conditions of a breeding chorus. Precisely how improvements in signal recognition and discrimination are brought about by the treefrogs' auditory system in the presence of spatially separated and temporally fluctuating noise is an area of ongoing investigation.

## 5. Future challenges

Continued progress will be made in understanding how evolution has solved cocktail-party-like problems by continuing to focus future work on the “right choice” of animal models. Treefrogs in the genus *Hyla* currently represent the most promising frog models for this line of research. While most work to date has been with three North American species, the genus itself is distributed across much of the globe throughout North and Central America, Europe, and Asia (Hua et al., 2009). Hence, researchers across different continents have relatively local access to the same genus of animal. Moreover, the vast amount of data amassed from behavioral, biophysical, physiological, and anatomical studies of hearing and sound communication in *Hyla*, combined with our understanding of evolutionary relationships in this genus (reviewed in Hua et al., 2009), surpasses that available for any other group of frogs and rivals what we know of other animal models in hearing research. This previous work provides a foundation of considerable breadth and depth for future studies. Treefrogs are also relatively easy to maintain in large numbers in laboratory environments. Because they are common pets, excellent guides to treefrog husbandry are available (Edmunds, 2007). Below I briefly outline some future directions for research on sound source perception in *Hyla*.

### 5.1 Single cues, multiple cues, and multiple sensory modalities

Thus far, most research on sound source perception in frogs has systematically investigated one acoustic cue at a time. This powerful experimental approach has a long history in psychoacoustics and in ethological studies of animal communication. The approach has so far identified roles for various spatial, spectral, and temporal cues in allowing frogs to perceive sound sources. However, we have only begun to scratch the surface when it comes to important questions about how these various cues function in auditory grouping. It remains to be determined, for example, whether differences in frequency on the order of among-individual differences promote auditory stream segregation in choruses of conspecifics. The extent to which harmonicity functions as an auditory grouping cue is presently unclear, and little work to date has asked whether common onsets and offsets also function in auditory grouping in frogs. Work on spatial proximity as a grouping cue has revealed surprising sensitivity for spatial coherence when it comes to simultaneous integration, but unexpected levels of tolerance for spatial incoherence when it comes to sequential integration. Can we explain these findings based on what we presently know about spatial hearing in treefrogs? Likewise, is it possible to predict based on the inherent

directionality of the frog's auditory system how much spatial release from masking should be expected?

Other open questions pertain to biological features of listeners. For instance, do different species, or the different sexes of a given species, assign different weights to the same cue? If so, why? Could species differences in auditory grouping be predicted by the acoustic structure of their signals or the chorus environments in which they communicate? Several species differences in scene analysis were noted in this review, and we know that even relatively closely related species can assign different weights to the same acoustic features used in call recognition (e.g., Schul and Bush, 2002). Identifying species differences in auditory grouping, if they exist, could provide a novel means to gaining additional insights into underlying mechanisms. Within a given species, it will be important to ask how tolerance for incoherence in a given cue relates to the perceptual limits for discriminating differences in that cue. New methodologies will be required to answer this question.

Still other unanswered (and in many way, un-asked) questions pertain to multiple cue use. Consequently, there is much to learn about how interactions between acoustic cues for grouping impact listener performance in source segregation tasks. To fully elucidate mechanisms, it will be necessary to investigate the extent to which variation in one cue modulates the use of other cues. Can frogs use multiple cues flexibly in ways that vary with the complexity of the acoustic environment? Does ambiguity in one cue cause frogs to shift weight to other cues? And what might be the importance of multiple sensory modalities? Recent studies of frogs using robotic frog models have demonstrated sensitivity to visual cues associated with calling, even in nocturnal species (e.g., Narins et al., 2003; Taylor et al., 2008; Taylor and Ryan, 2013). How might information gathered by the visual system interact with acoustic cues for auditory grouping to influence perceptual organization and communication (e.g., Talyor and Ryan, 2013)? The increasing use of robotics in animal behavior research will provide the means for answering this question.

## 5.2 Schema-based auditory scene analysis

As noted earlier, Bregman (1990) made an important distinction between “primitive” and “schema-based” mechanisms in auditory scene analysis. The latter depends on higher-level cognitive processes related to attention and learned expectations. While frogs must analyze acoustic scenes, they are not known within the animal kingdom for their cognitive prowess. Unlike, humans and songbirds, for example, frogs do not have to learn to produce or recognize their vocalizations (Gerhardt and Huber, 2002). Nevertheless, there would appear to be some scope for the operation of schema-based mechanisms in frogs. We know, for example, that male frogs in choruses often exhibit selective attention to the calls of their nearest neighbors so as to time their own calls in ways that reduce call overlap (reviewed in Schwartz and Bee, 2014). And while they do not learn their vocalizations, some frogs can learn to recognize specific individuals by associating individually distinct features of the caller's voice with a particular location in the chorus (Bee and Gerhardt, 2001a, b, 2002; Davis, 1987). In one frog with calls composed of a sequence of two distinct units produced in a fixed order, a listener's evolved template for recognizing properly ordered sequences has been suggested as an example of schema-based auditory grouping in action (Farris and

Ryan, 2011). Together, what these studies indicate is that auditory scene analysis in frogs cannot be comprehensively understood without considerations of potential schema-based processes. Thus, future studies aimed at understanding the roles of attention and either learned or evolved templates will be necessary to shed light on the importance and mechanisms of schema-based auditory scene analysis in frogs.

### 5.3 Energetic versus informational masking

The background noise typical of crowded social environments can produce two different types of masking, energetic masking and informational masking (reviewed in Durlach et al., 2003; Kidd et al., 2008; Watson, 2005). Energetic masking usually refers to masking caused by overlap of excitation patterns in the auditory periphery, though its effects might also be manifest at higher levels of auditory processing (Durlach et al., 2003). Informational masking – sometimes simply defined any masking that is not due to energetic masking – is thought to be related to auditory grouping and scene analysis, and to higher-level cognitive processes, such as learning and attention (Kidd et al., 2008). As such, it can involve multiple stages of auditory processing beyond the periphery (Durlach et al., 2003; Kidd et al., 2008). It should be noted, however, that such processing need not, by definition, relate only to complex cognitive processes to be susceptible to informational masking. We can put definitional issues aside and, in overly simplified terms, consider energetic masking as masking that makes a sound harder to hear (e.g., impaired detection), and informational masking as that which makes the sound harder to understand or describe (e.g., impaired feature analysis).

While the physiological basis of energetic masking is fairly well understood, this is not the case for informational masking. I believe future studies of treefrogs could shed some light on the physiology of informational masking. Research over the preceding decades has thoroughly described the “information” content of signals in terms of the acoustic features that are biologically relevant to decision making in these animals (reviewed in Gerhardt, 2001). We now also have a fairly sophisticated understanding of how the frog's auditory system becomes increasingly selective for behaviorally relevant (i.e., “informative”) features of signals at higher stages of processing (reviewed in Gerhardt and Huber, 2002; Hall, 1994; Narins et al., 2007; Wilczynski and Ryan, 2010). Preliminary evidence also suggests that informational masking occurs in treefrogs. For example, in the study of dip listening in *H. chrysoscelis* by Vélez and Bee (2011; see Fig. 11C), the nearly 6 dB increase in signal recognition thresholds in the presence of 40-Hz SAM noise, compared with a steady-state masker of equivalent RMS amplitude, can be interpreted as an instance of informational masking. The 40-Hz modulation in the masker presumably impaired central processing of the biologically important pulse rate of the stimulus (45.5 pulses/s). Additional studies explicitly targeting questions about informational masking in frogs would be welcomed. These studies could capitalize on what we already know about the processing of biologically relevant information by the frog's auditory periphery and brain to distinguish between the effects of energetic and informational masking at a physiological level.

## 5.4 Integrating psychophysics and physiology

Three methodological advances in particular would be helpful in advancing study of sound source perception in frogs. First, additional studies are needed that explicitly integrate behavioral experiments on sound source perception with biophysical and electrophysiological studies aimed at identifying the underlying mechanisms. In *Hyla*, the study of spatial proximity as a grouping cue in *H. versicolor* by Schwartz and Gerhardt (1995) represents an early example of just such a study. More such studies are needed if we are to make direct connections between hearing in real-world environments and its underlying mechanisms in frogs and other animal models. It will be important to integrate these approaches in future studies of the same species. Identifying interesting behaviors in one species (e.g., *Hyla chrysoscelis*) and describing its underlying mechanisms in another (e.g., *Rana temporaria*) that may be separated by tens or even hundreds of millions of years of evolution is best avoided when possible. Integrating psychophysical and physiological studies of the same awake, behaving individuals would, of course, be ideal.

Second, methods based on classical or operant conditioning would be useful for eliminating dependence on biologically meaningful signals. While the use of natural behavioral responses will remain a powerful approach in the future, alternative methods have potential to uncover perceptually salient features of stimuli that may have limited influence on phonotaxis and evoked calling, as demonstrated in studies of harmonicity in *H. cinerea* (Gerhardt et al., 1990; Simmons, 1988; Simmons et al., 1993). Previous efforts at conditioning frogs have met with mixed success (Simmons and Moss, 1995), but renewed efforts to overcome this methodological limitation might prove essential for making continued progress in unraveling the mechanisms of sound source perception in these animals. Advances along these lines would provide a means to investigate not only potential species differences in auditory scene analysis, but also potential sex differences, by eliminating any confounds stemming from differences in the experimental paradigm used (e.g., phonotaxis in females versus evoked calling in males).

Finally, rapid advances in neural imaging and recording techniques, such as two-photon imaging and optogenetics, are providing scientists with unprecedented opportunities to monitor and manipulate the activity of neural networks (e.g., Bathellier et al., 2012; Fenno et al., 2011; Grienberger et al., 2012; Stosiek et al., 2003). These techniques have not yet been used to study hearing and sound communication in frogs. Future efforts to develop these methods with frogs could significantly advance the integration of psychophysics and physiology in these animals. Fortunately, we have a wealth of previous knowledge on the anatomy and physiology of auditory processing in frogs to guide such efforts.

## 5.5 Quantitative modeling

Several studies have developed quantitative models of various aspects of auditory function in frogs, ranging from the physical response of the tympanum, to coding by auditory nerve fibers, to binaural interactions in the midbrain (Aertsen et al., 1986; Eggermont, 1985, 1993; Grigor'ev and Bibikov, 2010; Meenderink et al., 2005; Palmer and Pinder, 1984; Pinder and Palmer, 1983; Simmons et al., 1996; Van Stokkum, 1990; Van Stokkum and Gielen, 1989; Van Stokkum and Melssen, 1991; Wotton and Ferragamo, 2011). For the most part, these



models have been developed and parameterized using data from frogs in the genus *Rana*, such as *R. temporaria*, *R. catesbeiana*, and *R. esculenta*. In addition, these models have been used, primarily, to describe how the frog's auditory system extracts behaviorally salient features of communication signals in quiet. Less work has been done to evaluate how these models perform in the presence of biologically realistic noise. Therefore, I suggest two main avenues for future computational modeling efforts. First, the current but separate models for various features or processing stages of the frog auditory system should be better integrated into one computational model to the extent possible. New discoveries made since the original model formulations could be used to update and inform a unified model. One such direction would be to incorporate pulse interval counting neurons in the midbrain and combination sensitive neurons in the thalamus. To be widely useful to auditory researchers working with frogs, any such model would have to be constructed in such a way that the features most likely to vary among species (e.g., spectral sensitivity or temporal selectivity) could be easily parameterized and varied by different end users. Special efforts to adapt such a computational model to well-studied species of *Hyla* would be especially welcomed and could capitalize on the wealth of behavioral data already available for this genus. Second, such an integrated model should be used to understand and predict how the frog auditory system exploits the spatial, spectral, and temporal cues for auditory scene analysis discussed above. It will be necessary to shift our focus from one that uses models to understand how specific features of calls are encoded to one that predicts how the calls of individual conspecific males are perceived and compared amid high levels of background noise and acoustic clutter. Such a computational model of the frog auditory system could then be used to test hypotheses about the neural mechanisms of auditory scene analyses in these animals in ways that complement similar efforts based on models derived for humans (e.g., Mill et al., 2013; Shamma et al., 2011)

## 5.6 Summary

The cocktail party problem is a human problem in name only. It represents just one example of much more general problems of auditory scene analysis and acoustic communication that are encountered, and solved, by a diversity of other animals. Because of their ethological specializations for communicating in noise, and their demonstrated tractability in behavioral and physiological experiments, treefrogs in the genus *Hyla* provide one of the best animal models to study mechanisms for hearing in noisy social environments. As highlighted in this review, the treefrog auditory system solves a cocktail-party-like problem by exploiting some of the same acoustic cues that we humans also use to solve similar problems. Given the tinkering nature of evolution, important questions remain concerning precisely how they do so.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements

I thank Elyse Sussman and Mitchell Steinschneider for the invitation to contribute to this special issue; Michael Caldwell, Norman Lee, Katrina Schrode, Jessie Tanner, Jessica Ward, and two anonymous reviewers for helpful

feedback on previous versions of this manuscript; Michael Caldwell, Norman Lee, Vivek Nityananda, Katrina Schrode, Sandra Tekmen, Alejandro Vélez, and Jessica Ward for their hard work and dedication to much of the research described here; and the National Institute on Deafness and Other Communication Disorders (R03DC008396 and R01DC009582) and the National Science Foundation (IOS 0842759) for funding my research. All original research reported herein followed the “Principles of laboratory animal care” (NIH publication No. 86-23, revised 1985) and laws of the U.S.A., and was approved by the University of Minnesota’s Institutional Animal Care and Use Committee (#1202A10178).

## Abbreviations

<b>ABR</b>	auditory brainstem response
<b>AP</b>	amphibian papilla
<b>BP</b>	basilar papilla
<b>DPOAE</b>	Distortion product otoacoustic emission
<b>MRI</b>	magnetic resonance imaging
<b>RMS</b>	root mean square
<b>SNR</b>	signal-to-noise ratio
<b>SOAE</b>	Spontaneous otoacoustic emission
<b>SPL</b>	sound pressure level

## References

- Aertsen AMHJ, Vlaming MSMG, Eggermont JJ, Johannesma PIM. Directional hearing in the grassfrog (*Rana temporaria* L.). II. Acoustics and modelling of the auditory periphery. *Hear. Res.* 1986; 21:17–40. [PubMed: 3485627]
- Alder TB, Rose GJ. Long-term temporal integration in the anuran auditory system. *Nat. Neurosci.* 1998; 1:519–523. [PubMed: 10196551]
- Arak A. Sexual selection by male-male competition in natterjack toad choruses. *Nature.* 1983; 306:261–262.
- Arbogast TL, Mason CR, Kidd G. The effect of spatial separation on informational and energetic masking of speech. *J. Acoust. Soc. Am.* 2002; 112:2086–2098. [PubMed: 12430820]
- Aubin T, Jouventin P. Cocktail-party effect in king penguin colonies. *P. Roy. Soc. B-Biol. Sci.* 1998; 265:1665–1673.
- Aubin T, Jouventin P. How to vocally identify kin in a crowd: The penguin model. *Adv. Stud. Behav.* 2002; 31:243–277.
- Bacon SP, Grantham DW. Modulation masking: Effects of modulation frequency, depth, and phase. *J. Acoust. Soc. Am.* 1989; 85:2575–2580. [PubMed: 2745880]
- Bacon SP, Opie JM, Montoya DY. The effects of hearing loss and noise masking on the masking release for speech in temporally complex backgrounds. *Journal of Speech Language and Hear. Res.* 1998; 41:549–563.
- Bathellier B, Ushakova L, Rumpel S. Discrete neocortical dynamics predict behavioral categorization of sounds. *Neuron.* 2012; 76:435–449. [PubMed: 23083744]
- Bee MA. Sound source segregation in grey treefrogs: Spatial release from masking by the sound of a chorus. *Anim. Behav.* 2007; 74:549–558.
- Bee MA. Finding a mate at a cocktail party: spatial release from masking improves acoustic mate recognition in grey treefrogs. *Anim. Behav.* 2008a; 75:1781–1791. [PubMed: 19412318]
- Bee MA. Parallel female preferences for call duration in a diploid ancestor of an allotetraploid treefrog. *Anim. Behav.* 2008b; 76:845–853. [PubMed: 19727325]

- Bee MA. Spectral preferences and the role of spatial coherence in simultaneous integration in gray treefrogs (*Hyla chrysoscelis*). *J. Comp. Psychol.* 2010; 124:412–424. [PubMed: 20853948]
- Bee MA. Sound source perception in anuran amphibians. *Curr. Op. Neurobiol.* 2012; 22:301–310. [PubMed: 22265243]
- Bee MA, Gerhardt HC. Neighbour-stranger discrimination by territorial male bullfrogs (*Rana catesbeiana*): I. Acoustic basis. *Anim. Behav.* 2001a; 62:1129–1140.
- Bee MA, Gerhardt HC. Neighbour-stranger discrimination by territorial male bullfrogs (*Rana catesbeiana*): II. Perceptual basis. *Anim. Behav.* 2001b; 62:1141–1150.
- Bee MA, Gerhardt HC. Individual voice recognition in a territorial frog (*Rana catesbeiana*). *P. Roy. Soc. B-Biol. Sci.* 2002; 269:1443–1448.
- Bee MA, Micheyl C. The cocktail party problem: What is it? How can it be solved? And why should animal behaviorists study it? *J. Comp. Psychol.* 2008; 122:235–251. [PubMed: 18729652]
- Bee MA, Riemersma KK. Does common spatial origin promote the auditory grouping of temporally separated signal elements in grey treefrogs? *Anim. Behav.* 2008; 76:831–843. [PubMed: 19727419]
- Bee MA, Schwartz JJ. Behavioral measures of signal recognition thresholds in frogs in the presence and absence of chorus-shaped noise. *J. Acoust. Soc. Am.* 2009; 126:2788–2801. [PubMed: 19894854]
- Bee MA, Schwartz JJ, Summers K. All's well that begins Wells: Celebrating 60 years of *Anim. Behav.* and 36 years of research on anuran social behaviour. *Anim. Behav.* 2013; 85:5–18.
- Bee MA, Swanson EM. Auditory masking of anuran advertisement calls by road traffic noise. *Anim. Behav.* 2007; 74:1765–1776.
- Bee MA, Vélez A. Comodulation masking release in the perception of vocalizations by gray treefrogs. *Abstr. Assoc. Res. Otolaryngol.* 2008; 31:#812.
- Bee MA, Vélez A, Forester JD. Sound level discrimination by gray treefrogs in the presence and absence of chorus-shaped noise. *J. Acoust. Soc. Am.* 2012; 131:4188–4195. [PubMed: 22559390]
- Bernard, C. Introduction à l'étude de la médecine expérimentale. Baillière; Paris: 1865.
- Best V, Thompson ER, Mason CR, Kidd G. Spatial release from masking as a function of the spectral overlap of competing talkers. *J. Acoust. Soc. Am.* 2013; 133:3677–3680. [PubMed: 23742322]
- Bodnar DA. The separate and combined effects of harmonic structure, phase, and FM on female preferences in the barking treefrog (*Hyla gratiosa*). *J. Comp. Physiol. A.* 1996; 178:173–182. [PubMed: 8592302]
- Branstetter BK, Finneran JJ. Comodulation masking release in bottlenose dolphins (*Tursiops truncatus*). *J. Acoust. Soc. Am.* 2008; 124:625–633. [PubMed: 18647004]
- Branstetter BK, Trickey JS, Bakhtiari K, Black A, Aihara H, Finneran JJ. Auditory masking patterns in bottlenose dolphins (*Tursiops truncatus*) with natural, anthropogenic, and synthesized noise. *J. Acoust. Soc. Am.* 2013; 133:1811–1818. [PubMed: 23464049]
- Bregman, AS. Auditory Scene Analysis: The Perceptual Organization of Sound. MIT Press; Cambridge, MA: 1990.
- Bronkhorst AW. The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions. *Acustica.* 2000; 86:117–128.
- Bronkhorst AW, Plomp R. The effect of head-induced interaural time and level differences on speech intelligibility in noise. *J. Acoust. Soc. Am.* 1988; 83:1508–1516. [PubMed: 3372866]
- Bronkhorst AW, Plomp R. Effect of multiple speech-like maskers on binaural speech recognition in normal and impaired hearing. *J. Acoust. Soc. Am.* 1992; 92:3132–3139. [PubMed: 1474228]
- Buerkle NP, Schrode KM, Bee MA. The auditory brainstem response in green treefrogs (*Hyla cinerea*). *Comp. Biochem. Physiol. A.* in review.
- Burmeister SS, Mangiamela LA, Lebonville CL. Acoustic modulation of immediate early gene expression in the auditory midbrain of female tungara frogs. *Brain Res.* 2008; 1190:105–114. [PubMed: 18061149]
- Caldwell MS, Bee MA. Spatial hearing in Cope's gray treefrog: I. Open and closed loop experiments on sound localization in the presence and absence of noise. *J. Comp. Physiol. A.* submitted.

- Caldwell MS, Lee N, Johns AR, Christensen-Dalsgaard J, Bee MA. Spatial hearing in Cope's gray treefrog: II. Frequency-dependent directionality of the amplitude and phase of tympanum vibrations. *J. Comp. Physiol. A*. submitted.
- Cannatella, D. [November 09, 2013] *Vieraella herbsti*. The Oldest True Frog. Version 01 January 1995.. 1995. [http://tolweb.org/Vieraella\\_herbsti/16964/1995.01.01](http://tolweb.org/Vieraella_herbsti/16964/1995.01.01) in The Tree of Life Web Project, <http://tolweb.org/>
- Capranica, RR. The Evoked Vocal Response of the Bullfrog: A Study of Communication by Sound. M.I.T. Press; Cambridge, MA: 1965.
- Capranica, RR. Morphology and physiology of the auditory system. In: Llinas, R.; Precht, W., editors. *Frog Neurobiology: A Handbook*. Springer-Verlag; New York: 1976. p. 551-575.
- Capranica, RR.; Moffat, JM. Neurobehavioral correlates of sound communication in anurans. In: Ewert, JP.; Capranica, RR.; Ingle, DJ., editors. *Advances in Vertebrate Neuroethology*. Plenum Press; New York: 1983. p. 701-730.
- Carlyon RP. How the brain separates sounds. *Trends Cog. Sci.* 2004; 8:465–471.
- Carlyon, RP.; Gockel, H. Effects of harmonicity and regularity on the perception of sound sources. In: Yost, WA.; Popper, AN.; Fay, RR., editors. *Auditory Perception of Sound Sources*. Springer; New York: 2008. p. 191-213.
- Cherry EC. Some experiments on the recognition of speech, with one and with two ears. *J. Acoust. Soc. Am.* 1953; 25:975–979.
- Christensen-Dalsgaard, J. Directional hearing in nonmammalian tetrapods. In: Popper, AN.; Fay, RR., editors. *Sound Source Localization*. Springer; New York: 2005. p. 67-123.
- Christensen-Dalsgaard J. Vertebrate pressure-gradient receivers. *Hear. Res.* 2011; 273:37–45. [PubMed: 20727396]
- Christensen-Dalsgaard J, Carr CE. Evolution of a sensory novelty: Tympanic ears and the associated neural processing. *Brain Res. Bull.* 2008; 75:365–370. [PubMed: 18331899]
- Cooke M, Hershey JR, Rennie SJ. Monaural speech separation and recognition challenge. *Comput. Speech Lang.* 2010; 24:1–15.
- Darwin, CJ. Pitch and Auditory Grouping. In: Plack, CJ.; Oxenham, AJ.; Fay, RR.; Popper, AN., editors. *Pitch: Neural Coding and Perception*. Springer; New York: 2005. p. 278-305.
- Darwin CJ. Listening to speech in the presence of other sounds. *Philos. T. Roy. Soc. B.* 2008a; 363:1011–1021.
- Darwin, CJ. Spatial hearing and perceiving sources. In: Yost, WA.; Popper, AN.; Fay, RR., editors. *Auditory Perception of Sound Sources*. Springer; New York: 2008b. p. 215-232.
- Darwin, CJ.; Carlyon, RP. Auditory grouping. In: Moore, BCJ., editor. *Hearing*. Academic Press; New York: 1995. p. 387-424.
- Davis MS. Acoustically mediated neighbor recognition in the North American bullfrog, *Rana catesbeiana*. *Behav. Ecol. Sociobiol.* 1987; 21:185–190.
- Diekamp B, Gerhardt HC. Selective phonotaxis to advertisement calls in the gray treefrog *Hyla versicolor*: Behavioral experiments and neurophysiological correlates. *J. Comp. Physiol. A.* 1995; 177:173–190. [PubMed: 7636766]
- Dirks DD, Wilson RH. The effect of spatially separated sound sources on speech intelligibility. *Journal of Speech and Hear. Res.* 1969; 12:5–38. [PubMed: 5779912]
- Durlach NI, Mason CR, Kidd G, Arbogast TL, Colburn HS, Shinn-Cunningham BG. Note on informational masking (L). *J. Acoust. Soc. Am.* 2003; 113:2984–2987. [PubMed: 12822768]
- Edmunds, D. *Tree Frogs: A Complete Guide to Arboreal Frogs*. T.F.H. Publications; Neptune City, NJ: 2007.
- Edwards CJ, Alder TB, Rose GJ. Auditory midbrain neurons that count. *Nat. Neurosci.* 2002; 5:934–936. [PubMed: 12219094]
- Eggermont JJ. Peripheral auditory adaptation and fatigue: A model oriented review. *Hear. Res.* 1985; 18:57–71. [PubMed: 2993220]
- Eggermont JJ. Wiener and Volterra analyses applied to the auditory system. *Hear. Res.* 1993; 66:177–201. [PubMed: 8509309]

- Ehret G, Capranica RR. Masking patterns and filter characteristics of auditory nerve fibers in the green treefrog (*Hyla cinerea*). *J. Comp. Physiol.* 1980; 141:1–12.
- Ehret G, Gerhardt HC. Auditory masking and effects of noise on responses of the green treefrog (*Hyla cinerea*) to synthetic mating calls. *J. Comp. Physiol. A.* 1980; 141:13–18.
- Ehret G, Moffat AJM, Capranica RR. Two-tone suppression in auditory nerve fibers of the green treefrog (*Hyla cinerea*). *J. Acoust. Soc. Am.* 1983; 73:2093–2095. [PubMed: 6875094]
- Elliott TM, Christensen-Dalsgaard J, Kelley DB. Temporally selective processing of communication signals by auditory midbrain neurons. *J. Neurophysiol.* 2011; 105:1620–1632. [PubMed: 21289132]
- Farris HE, Rand AS, Ryan MJ. The effects of spatially separated call components on phonotaxis in túngara frogs: Evidence for auditory grouping. *Brain Behav. Evolut.* 2002; 60:181–188.
- Farris HE, Rand AS, Ryan MJ. The effects of time, space and spectrum on auditory grouping in túngara frogs. *J. Comp. Physiol. A.* 2005; 191:1173–1183.
- Farris HE, Ryan MJ. Relative comparisons of call parameters enable auditory grouping in frogs. *Nat. Comm.* 2011; 2:410.
- Fay RR. Signal-to-noise ratio for source determination and for a comodulated masker in goldfish, *Carassius auratus*. *J. Acoust. Soc. Am.* 2011; 129:3367–3372. [PubMed: 21568437]
- Fay RR, Popper AN. Evolution of hearing in vertebrates: The inner ears and processing. *Hear. Res.* 2000; 149:1–10. [PubMed: 11033242]
- Fay, RR.; Simmons, AM. The sense of hearing in fishes and amphibians. In: Fay, RR.; Popper, AN., editors. *Comparative Hearing: Fish and Amphibians*. Springer; New York: 1999. p. 269-318.
- Fellers GM. Aggression, territoriality, and mating behaviour in North American treefrogs. *Anim. Behav.* 1979; 27:107–119.
- Feng AS, Capranica RR. Sound localization in anurans II. Binaural interaction in superior olivary nucleus of the green tree frog (*Hyla cinerea*). *J. Neurophysiol.* 1978; 41:43–54. [PubMed: 621545]
- Feng, AS.; Schul, J. Sound processing in real-world environments. In: Narins, PA.; Feng, AS.; Fay, RR.; Popper, AN., editors. *Hearing and Sound Communication in Amphibians*. Springer; New York: 2007. p. 323-350.
- Fenno L, Yizhar O, Deisseroth K, Hyman SE, Jessell TM, Shatz CJ, Stevens CF, Zoghbi HY. The development and application of optogenetics. *Annu. Rev. Neurosci.* 2011; 34:389–412. [PubMed: 21692661]
- Festen JM, Plomp R. Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *J. Acoust. Soc. Am.* 1990; 88:1725–1736. [PubMed: 2262629]
- Freyman RL, Helfer KS, McCall DD, Clifton RK. The role of perceived spatial separation in the unmasking of speech. *J. Acoust. Soc. Am.* 1999; 106:3578–3588. [PubMed: 10615698]
- Frishkopf LS, Goldstein Jr MH. Responses to acoustic stimuli from single units in the eighth nerve of the bullfrog. *The J. Acoust. Soc. Am.* 1963; 35:1219–1228.
- Fritzsche, B.; Wolkowiak, W.; Ryan, MJ.; Wilczynski, W.; Hetherington, T. *The Evolution of the Amphibian Auditory System*. Wiley; New York: 1988.
- Füllgrabe C, Berthommier F, Lorenzi C. Masking release for consonant features in temporally fluctuating background noise. *Hear. Res.* 2006; 211:74–84. [PubMed: 16289579]
- Gerhardt HC. Sound pressure levels and radiation patterns of vocalizations of some North American frogs and toads. *J. Comp. Physiol.* 1975; 102:1–12.
- Gerhardt HC. Evolutionary and neurobiological implications of selective phonotaxis in the green treefrog, *Hyla cinerea*. *Anim. Behav.* 1987; 35:1479–1489.
- Gerhardt, HC. Acoustic properties used in call recognition by frogs and toads. In: Fritzsche, B.; Ryan, MJ.; Wilczynski, W.; Hetherington, TE.; Wolkowiak, W., editors. *The Evolution of the Amphibian Auditory System*. Wiley & Sons; New York: 1988. p. 455-483.
- Gerhardt, HC. Phonotaxis in female frogs and toads: Execution and design of experiments. In: Klump, GM.; Dooling, RJ.; Fay, RR.; Stebbins, WC., editors. *Methods in Comparative Psychoacoustics*. Birkhäuser Verlag; Basel: 1995. p. 209-220.

- Gerhardt HC. Acoustic communication in two groups of closely related treefrogs. *Adv. Stud. Behav.* 2001; 30:99–167.
- Gerhardt HC. Acoustic spectral preferences in two cryptic species of grey treefrogs: Implications for mate choice and sensory mechanisms. *Anim. Behav.* 2005; 70:39–48.
- Gerhardt HC, Allan S, Schwartz JJ. Female green treefrogs (*Hyla cinerea*) do not selectively respond to signals with a harmonic structure in noise. *J. Comp. Physiol. A.* 1990; 166:791–794. [PubMed: 2359052]
- Gerhardt, HC.; Bee, MA. Recognition and localization of acoustic signals. In: Narins, PM.; Feng, AS.; Fay, RR.; Popper, AN., editors. *Hearing and Sound Communication in Amphibians*. Springer; New York: 2007. p. 113-146.
- Gerhardt HC, Doherty JA. Acoustic communication in the gray treefrog, *Hyla versicolor*: Evolutionary and neurobiological implications. *J. Comp. Physiol. A.* 1988; 162:261–278.
- Gerhardt HC, Dyson ML, Tanner SD. Dynamic properties of the advertisement calls of gray tree frogs: Patterns of variability and female choice. *Behav. Ecol.* 1996; 7:7–18.
- Gerhardt HC, Huber F. *Acoustic Communication in Insects and Anurans: Common Problems and Diverse Solutions*. 2002:531.
- Gerhardt HC, Klump GM. Masking of acoustic signals by the chorus background noise in the green treefrog: A limitation on mate choice. *Anim. Behav.* 1988; 36:1247–1249.
- Gerhardt HC, Martinez-Rivera CC, Schwartz JJ, Marshall VT, Murphy CG. Preferences based on spectral differences in acoustic signals in four species of treefrogs (Anura: Hylidae). *J. Exp. Bio.* 2007; 210:2990–2998. [PubMed: 17704074]
- Gerhardt HC, Ptacek MB, Barnett L, Torke KG. Hybridization in the diploid-tetraploid treefrogs *Hyla chrysoscelis* and *Hyla versicolor*. *Copeia.* 1994; 1994:51–59.
- Gerhardt HC, Rheinlaender J. Localization of an elevated sound source by the green tree frog. *Science.* 1982; 217:663–664. [PubMed: 17817539]
- Gerhardt HC, Roberts JD, Bee MA, Schwartz JJ. Call matching in the quacking frog (*Crinia georgiana*). *Behav. Ecol. Sociobiol.* 2000a; 48:243–251.
- Gerhardt HC, Schul J. A quantitative analysis of behavioral selectivity for pulse rise-time in the gray treefrog, *Hyla versicolor*. *J. Comp. Physiol. A.* 1999; 185:33–40. [PubMed: 10450610]
- Gerhardt HC, Tanner SD, Corrigan CM, Walton HC. Female preference functions based on call duration in the gray tree frog (*Hyla versicolor*). *Behav. Ecol.* 2000b; 11:663–669.
- Goense JBM, Feng AS. Effects of noise bandwidth and amplitude modulation on masking in frog auditory midbrain neurons. *PLoS ONE.* 2012; 7:e31589. [PubMed: 22348114]
- Grienberger C, Adelsberger H, Stroh A, Milos RI, Garaschuk O, Schierloh A, Nelken I, Konnerth A. Sound-evoked network calcium transients in mouse auditory cortex in vivo. *J. Physiol.-London.* 2012; 590:899–918. [PubMed: 22106174]
- Grigor'ev DY, Bibikov NG. Model of a neuron trained to extract periodicity. *Acoust. Phys.* 2010; 56:720–728.
- Gustafsson HA, Arlinger SD. Masking of speech by amplitude-modulated noise. *J. Acoust. Soc. Am.* 1994; 95:518–529. [PubMed: 8120263]
- Hall JC. Central processing of communication sounds in the anuran auditory system. *Am. Zool.* 1994; 34:670–684.
- Hawley ML, Litovsky RY, Colburn HS. Speech intelligibility and localization in a multi-source environment. *J. Acoust. Soc. Am.* 1999; 105:3436–3448. [PubMed: 10380667]
- Hedges SB, Dudley J, Kumar S. TimeTree: A public knowledge-base of divergence times among organisms. *Bioinformatics.* 2006; 22:2971–2972. [PubMed: 17021158]
- Hillery CM. Seasonality of two midbrain auditory responses in the treefrog, *Hyla chrysoscelis*. *Copeia.* 1984; 1984:844–852.
- Hillery CM, Fay RR. Forward masking and suppression in the midbrain of the southern grey treefrog (*Hyla chrysoscelis*). *J. Comp. Physiol.* 1982; 146:435–447.
- Hirsh IJ. The relation between localization and intelligibility. *J. Acoust. Soc. Am.* 1950; 22:196–200.
- Ho CCK, Narins PM. Directionality of the pressure-difference receiver ears in the northern leopard frog, *Rana pipiens pipiens*. *J. Comp. Physiol. A.* 2006; 192:417–429.



- Hoke KL, Burmeister SS, Fernald RD, Rand AS, Ryan MJ, Wilczynski W. Functional mapping of the auditory midbrain during mate call reception. *J. Neurosci.* 2004; 24:11264–11272. [PubMed: 15601932]
- Holloway AK, Cannatella DC, Gerhardt HC, Hillis DM. Polyploids with different origins and ancestors form a single sexual polyploid species. *Am. Nat.* 2006; 167:E88–E101. [PubMed: 16670990]
- Houtgast T, Steeneken HJM. A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria. *J. Acoust. Soc. Am.* 1985; 77:1069–1077.
- Hua X, Fu CZ, Li JT, de Oca ANM, Wiens JJ. A revised phylogeny of holarctic treefrogs (genus *Hyla*) based on nuclear and mitochondrial DNA sequences. *Herpetologica.* 2009; 65:246–259.
- Hulse SH. Auditory scene analysis in animal communication. *Adv. Stud. Behav.* 2002; 31:163–200.
- Jacob F. Evolution and tinkering. *Science.* 1977; 196:1161–1166. [PubMed: 860134]
- Jensen KK. Comodulation detection differences in the hooded crow (*Corvus corone cornix*), with direct comparison to human subjects. *J. Acoust. Soc. Am.* 2007; 121:1783–1789. [PubMed: 17407915]
- Jørgensen MB. Comparative studies of the biophysics of directional hearing in anurans. *J. Comp. Physiol. A.* 1991; 169:591–598.
- Jørgensen MB, Gerhardt HC. Directional hearing in the gray tree frog *Hyla versicolor*: Eardrum vibrations and phonotaxis. *J. Comp. Physiol. A.* 1991; 169:177–183. [PubMed: 1748974]
- Kashino M. Phonemic restoration: The brain creates missing speech sounds. *Acoustical Science and Technology.* 2006; 27:318.
- Kelley DB. Vocal communication in frogs. *Curr. Op. Neurobiol.* 2004; 14:751–757. [PubMed: 15582379]
- Kidd, G.; Mason, CR.; Richards, VM.; Gallun, FJ.; Durlach, NI. Informational masking. In: Yost, WA.; Popper, AN.; Fay, RR., editors. *Auditory Perception of Sound Sources.* Springer; New York: 2008. p. 143-189.
- King AJ. Auditory neuroscience: Filling in the gaps. *Curr. Biol.* 2007; 17:R799–R801. [PubMed: 17878047]
- Klump, GM. Bird communication in the noisy world. In: Kroodsma, DE.; Miller, EH., editors. *Ecology and Evolution of Acoustic Communication in Birds.* Cornell University Press; Ithaca, NY, pp: 1996. p. 321-338.
- Klump GM, Benedix JH, Gerhardt HC, Narins PM. AM representation in green treefrog auditory nerve fibers: Neuroethological implications for pattern recognition and sound localization. *J. Comp. Physiol. A.* 2004; 190:1011–1021.
- Klump GM, Gerhardt HC. Sound localization in the barking treefrog. *Naturwissenschaften.* 1989; 76:35–37. [PubMed: 2704420]
- Klump GM, Langemann U. Comodulation masking release in a songbird. *Hear. Res.* 1995; 87:157–164. [PubMed: 8567432]
- Konishi M. Coding of auditory space. *Annu. Rev. Neurosci.* 2003; 26:31–55. [PubMed: 14527266]
- Krebs HA. The August Krogh Principle: “For many problems there is an animal on which it can be most conveniently studied”. *J. Exp. Zool.* 1975; 194:221–226. [PubMed: 811756]
- Krogh A. The progress of physiology. *Am. J. Physiol.* 1929; 90:243–251.
- Kwon BJ, Turner CW. Consonant identification under maskers with sinusoidal modulation: Masking release or modulation interference? *J. Acoust. Soc. Am.* 2001; 110:1130–1140. [PubMed: 11519580]
- Langemann U, Klump GM. Detecting modulated signals in modulated noise: I. Behavioural auditory thresholds in a songbird. *Eur. J. Neurosci.* 2007; 26:1969–1978. [PubMed: 17897400]
- Lee N, Micheyl C, Vélez A, Bee MA. Comodulation masking release (CMR) in treefrogs and a frog-inspired auditory filterbank to investigate its neural correlates. *Abstr. Assoc. Res. Otolaryngol.* 2013; 36:#209.
- Lim D, Capranica RR. Measurement of temporal regularity of spike train responses in auditory nerve fibers of the green treefrog. *J. Neurosci. Meth.* 1994; 52:203–213.

- Lin WY, Feng AS. Free-field unmasking response characteristics of frog auditory nerve fibers: Comparison with the responses of midbrain auditory neurons. *J. Comp. Physiol. A.* 2001; 187:699–712. [PubMed: 11778832]
- Lin WY, Feng AS. GABA is involved in spatial unmasking in the frog auditory midbrain. *J. Neurosci.* 2003; 23:8143–8151. [PubMed: 12954877]
- Lombard RE, Straughan IR. Functional aspects of anuran middle ear structures. *J. Exp. Bio.* 1974; 61:71–93. [PubMed: 4606567]
- Manley, GA.; Popper, AN.; Fay, RR. *Evolution of the Vertebrate Auditory System.* Springer; New York: 2004.
- Marshall VT, Schwartz JJ, Gerhardt HC. Effects of heterospecific call overlap on the phonotactic behaviour of grey treefrogs. *Anim. Behav.* 2006; 72:449–459.
- Mason, MJ. Pathways for sound transmission to the inner ear in amphibians. In: Narins, PM.; Feng, AS.; Fay, RR.; Popper, AN., editors. *Hearing and Sound Communication in Amphibians.* Springer; New York: 2007. p. 147-183.
- McDermott JH. The cocktail party problem. *Curr. Biol.* 2009; 19:R1024–R1027. [PubMed: 19948136]
- Meenderink SWF, van Dijk P, Narins PM. Comparison between distortion product otoacoustic emissions and nerve fiber responses from the basilar papilla of the frog. *J. Acoust. Soc. Am.* 2005; 117:3165–3173. [PubMed: 15957784]
- Megela-Simmons A, Moss CF, Daniel KM. Behavioral audiograms of the bullfrog (*Rana catesbeiana*) and the green tree frog (*Hyla cinera*). *J. Acoust. Soc. Am.* 1985; 78:1236–1244. [PubMed: 3877086]
- Michelsen A, Jørgensen MB, Christensen-Dalsgaard J, Capranica RR. Directional hearing of awake, unrestrained treefrogs *Naturwissenschaften.* 1986; 73:682–683.
- Miles RN, Hoy RR. The development of a biologically-inspired directional microphone for hearing aids. *Audio. Neuro-Otol.* 2006; 11:86–94.
- Mill RW, Bohm TM, Bendixen A, Winkler I, Denham SL. Modelling the emergence and dynamics of perceptual organisation in auditory streaming. *PLoS Comput. Biol.* 2013;9.
- Miller G. Behavioral neuroscience uncaged. *Science.* 2004; 306:432–434. [PubMed: 15486287]
- Miranda JA, Wilczynski W. Female reproductive state influences the auditory midbrain response. *J. Comp. Physiol. A.* 2009a; 195:341–349.
- Miranda JA, Wilczynski W. Sex differences and androgen influences on midbrain auditory thresholds in the green treefrog, *Hyla cinerea*. *Hear. Res.* 2009b; 252:79–88. [PubMed: 19371774]
- Moore BCJ. Speech processing for the hearing-impaired: Successes, failures, and implications for speech mechanisms. *Speech Comm.* 2003; 41:81–91.
- Moore BCJ, Gockel H. Factors influencing sequential stream segregation. *Acta Acust. United Ac.* 2002; 88:320–333.
- Moore BCJ, Gockel HE. Properties of auditory stream formation. *Philos. T. Roy. Soc. B.* 2012; 367:919–931.
- Moss CF, Simmons AM. Frequency selectivity of hearing in the green treefrog, *Hyla cinerea*. *J. Comp. Physiol. A.* 1986; 159:257–266. [PubMed: 3761225]
- Mudry KM, Capranica RR. Correlation between auditory thalamic area evoked responses and species-specific call characteristics II. *Hyla cinerea* (Anura: Hylidae). *J. Comp. Physiol. A.* 1987; 161:407–416. [PubMed: 3668880]
- Narins PM. Effects of masking noise on evoked calling in the Puerto Rican coqui (Anura, Leptodactylidae). *J. Comp. Physiol.* 1982; 147:439–446.
- Narins, PM.; Feng, AS.; Fay, RR.; Popper, AN. Springer; New York: 2007. *Hearing and Sound Communication in Amphibians.*; p. 362
- Narins PM, Hödl W, Grabul DS. Bimodal signal requisite for agonistic behavior in a dart-poison frog, *Epiplatys femoralis*. *P. Natl. Acad. Sci. USA.* 2003; 100:577–580.
- Narins, PM.; Zelick, R. The effects of noise on auditory processing and behavior in amphibians. In: Fritsch, B.; Ryan, MJ.; Wilczynski, W.; Hetherington, TE.; Walkowiak, W., editors. *The Evolution of the Amphibian Auditory System.* Wiley & Sons; New York: 1988. p. 511-536.

- Nelken I, Rotman Y, Bar Yosef O. Responses of auditory-cortex neurons to structural features of natural sounds. *Nature*. 1999; 397:154–157. [PubMed: 9923676]
- Nityananda V, Bee MA. Finding your mate at a cocktail party: Frequency separation promotes auditory stream segregation of concurrent voices in multi-species frog choruses. *PLoS ONE*. 2011; 6:e21191. [PubMed: 21698268]
- Nityananda V, Bee MA. Spatial release from masking in a free-field source identification task by gray treefrogs. *Hear. Res.* 2012; 285:86–97. [PubMed: 22240459]
- Odom, SE.; Leary, CJ.; Graham, JA.; Vasquez-Opazo, GA.; Rose, GJ. Physiological and anatomical differentiation of the inferior colliculus in the two species of gray treefrogs, *H. chrysoscelis* and *H. versicolor*. Program No. 502.09. 2012 Neuroscience Meeting Planner.. Society for Neuroscience, 2012; New Orleans, LA: 2012. Online
- Palmer AR, Pinder AC. The directionality of the frog ear described by a mechanical model. *J. Theor. Biol.* 1984; 110:205–215. [PubMed: 6492830]
- Peissig J, Kollmeier B. Directivity of binaural noise reduction in spatial multiple noise-source arrangements for normal and impaired listeners. *J. Acoust. Soc. Am.* 1997; 101:1660–1670. [PubMed: 9069633]
- Penna M, Capranica RR, Somers J. Hormone-induced vocal behavior and midbrain auditory sensitivity in the green treefrog, *Hyla cinerea*. *J. Comp. Physiol. A.* 1992; 170:73–82. [PubMed: 1573572]
- Pinder AC, Palmer AR. Mechanical properties of the frog ear: Vibration measurements under free- and closed-field acoustic conditions. *P. Roy. Soc. B-Biol. Sci.* 1983; 219:371–396.
- Plomp R, Mimpen AM. Effect of the orientation of the speaker's head and azimuth of a noise source on the speech reception threshold for sentences. *Acustica.* 1981; 48:325–328.
- Popper AN, Fay RR. Evolution of the ear and hearing: Issues and questions. *Brain Behav. Evolut.* 1997; 50:213–221.
- Ptacek MB, Gerhardt HC, Sage RD. Speciation by polyploidy in treefrogs: Multiple origins of the tetraploid, *Hyla versicolor*. *Evolution.* 1994; 48:898–908.
- Pyron RA, Wiens JJ. A large-scale phylogeny of Amphibia including over 2800 species, and a revised classification of extant frogs, salamanders, and caecilians. *Mol. Phylogenet. Evol.* 2011; 61:543–583. [PubMed: 21723399]
- Ratnam R, Feng AS. Detection of auditory signals by frog inferior collicular neurons in the presence of spatially separated noise. *J. Neurophysiol.* 1998; 80:2848–2859. [PubMed: 9862889]
- Rheinlaender J, Gerhardt HC, Yager DD, Capranica RR. Accuracy of phonotaxis by the green treefrog (*Hyla cinerea*). *J. Comp. Physiol.* 1979; 133:247–255.
- Rheinlaender, J.; Klump, GM. Behavioral aspects of sound localization. In: Fritzsche, B.; Ryan, MJ.; Wilczynski, W.; Hetherington, T., editors. *The Evolution of the Amphibian Auditory System*. Wiley & Sons; New York: 1988. p. 297-305.
- Rheinlaender J, Walkowiak W, Gerhardt HC. Directional hearing in the green treefrog: A variable mechanism? *Naturwissenschaften.* 1981; 68:430–431.
- Richards DG, Wiley RH. Reverberations and amplitude fluctuations in the propagation of sound in a forest: Implications for animal communication. *Am. Nat.* 1980; 115:381–399.
- Römer, H. Masking by noise in acoustic insects: Problems and solutions. In: Brumm, H., editor. *Animal Communication and Noise*. Springer; New York: 2014. in press
- Rose GJ, Brenowitz EA, Capranica RR. Species specificity and temperature dependency of temporal processing by the auditory midbrain of two species of treefrogs. *J. Comp. Physiol. A.* 1985; 157:763–769. [PubMed: 3837112]
- Rose GJ, Capranica RR. Processing amplitude-modulated sounds by the auditory midbrain of two species of toads: Matched temporal filters. *J. Comp. Physiol.* 1984; 154:211–219.
- Rose GJ, Leary CJ, Edwards CJ. Interval-counting neurons in the anuran auditory midbrain: Factors underlying diversity of interval tuning. *J. Comp. Physiol. A.* 2011; 197:97–108.
- Ryan, MJ. *Anuran Communication*. Smithsonian Institution Press; Washington D.C.: 2001. p. 252
- Schnupp JWH, Carr CE. On hearing with more than one ear: Lessons from evolution. *Nat. Neurosci.* 2009; 12:692–697. [PubMed: 19471267]

- Schrode K, Ward JL, Vélez A, Bee MA. Female preferences for spectral call properties in the western genetic lineage of Cope's gray treefrog (*Hyla chrysoscelis*). *Behav. Ecol. Sociobiol.* 2012; 66:1595–1606. [PubMed: 24496093]
- Schrode KM, Buerkle NP, Brittan-Powell EF, Bee MA. Auditory brainstem responses in Cope's gray treefrog (*Hyla chrysoscelis*): Effects of frequency, level, sex and size. *J. Comp. Physiol. A.* in review.
- Schul J, Bush SL. Non-parallel coevolution of sender and receiver in the acoustic communication system of treefrogs. *P. Roy. Soc. B-Biol. Sci.* 2002; 269:1847–1852.
- Schwartz JJ. The function of call alternation in anuran amphibians: A test of three hypotheses. *Evolution.* 1987; 41:461–471.
- Schwartz JJ.; Bee, MA. Anuran acoustic signal production in noisy environments. In: Brumm, H., editor. *Animal Communication and Noise*. Springer; New York: 2014. in press
- Schwartz JJ, Brown R, Turner S, Dushaj K, Castano M. Interference risk and the function of dynamic shifts in calling in the gray treefrog (*Hyla versicolor*). *J. Comp. Psychol.* 2008; 122:283–288. [PubMed: 18729656]
- Schwartz JJ, Buchanan BW, Gerhardt HC. Female mate choice in the gray treefrog (*Hyla versicolor*) in three experimental environments. *Behav. Ecol. Sociobiol.* 2001; 49:443–455.
- Schwartz JJ, Crimarco NC, Bregman Y, Umeoji K. An investigation of the functional significance of responses of the gray treefrog (*Hyla versicolor*) to chorus noise. *J. Herpetol.* 2013; 47:354–360.
- Schwartz JJ, Gerhardt HC. Spatially mediated release from auditory masking in an anuran amphibian. *J. Comp. Physiol. A.* 1989; 166:37–41.
- Schwartz JJ, Gerhardt HC. Directionality of the auditory system and call pattern recognition during acoustic interference in the gray treefrog, *Hyla versicolor*. *Audit. Neurosci.* 1995; 1:195–206.
- Schwartz JJ, Huth K, Jones SH, Brown R, Marks J. Tests for call restoration in the gray treefrog, *Hyla versicolor*. *Bioacoustics.* 2010; 20:59–86.
- Schwartz JJ, Simmons AM. Encoding of a spectrally complex natural call in the bullfrog's auditory nerve. *J. Comp. Physiol. A.* 1990; 166:489–499. [PubMed: 2332839]
- Seeba F, Schwartz JJ, Bee MA. Testing an auditory illusion in frogs: Perceptual restoration or sensory bias? *Anim. Behav.* 2010; 79:1317–1328. [PubMed: 20514342]
- Seok J, Warren HS, Cuenca AG, Mindrinos MN, Baker HV, Xu W, Richards DR, McDonald-Smith GP, Gao H, Hennessy L, Finnerty CC, Lopez CM, Honari S, Moore EE, Minei JP, Cuschieri J, Bankey PE, Johnson JL, Sperry J, Nathens AB, Billiar TR, West MA, Jeschke MG, Klein MB, Gamelli RL, Gibran NS, Brownstein BH, Miller-Graziano C, Calvano SE, Mason PH, Cobb JP, Rahme LG, Lowry SF, Maier RV, Moldawer LL, Herndon DN, Davis RW, Xiao W, Tompkins RG. Genomic responses in mouse models poorly mimic human inflammatory diseases. *Proc. Natl. Acad. Sci USA.* 2013; 110:3507–3512. [PubMed: 23401516]
- Shamma SA, Elhilali M, Micheyl C. Temporal coherence and attention in auditory scene analysis. *Trends Neurosci.* 2011; 34:114–123. [PubMed: 21196054]
- Shinn-Cunningham BG, Ihlefeld A, Satyavarta Larson, E. Bottom-up and top-down influences on spatial unmasking. *Acta Acust. United Ac.* 2005; 91:967–979.
- Shinn-Cunningham BG, Schickler J, Kopco N, Litovsky R. Spatial unmasking of nearby speech sources in a simulated anechoic environment. *J. Acoust. Soc. Am.* 2001; 110:1118–1129. [PubMed: 11519579]
- Simmons AM. Selectivity for harmonic structure in complex sounds by the green treefrog (*Hyla cinerea*). *J. Comp. Physiol. A.* 1988; 162:397–403. [PubMed: 3351790]
- Simmons AM. “To Ear is Human, to Frogive is Divine”: Bob Capranica's legacy to auditory neuroethology. *J. Comp. Physiol. A.* 2013; 199:169–182.
- Simmons AM, Bean ME. Perception of mistuned harmonics in complex sounds by the bullfrog (*Rana catesbeiana*). *J. Comp. Psychol.* 2000; 114:167–173. [PubMed: 10890588]
- Simmons AM, Buxbaum RC, Mirin MP. Perception of complex sounds by the green treefrog, *Hyla cinerea*: Envelope and fine-structure cues. *J. Comp. Physiol. A.* 1993; 173:321–327. [PubMed: 8229896]

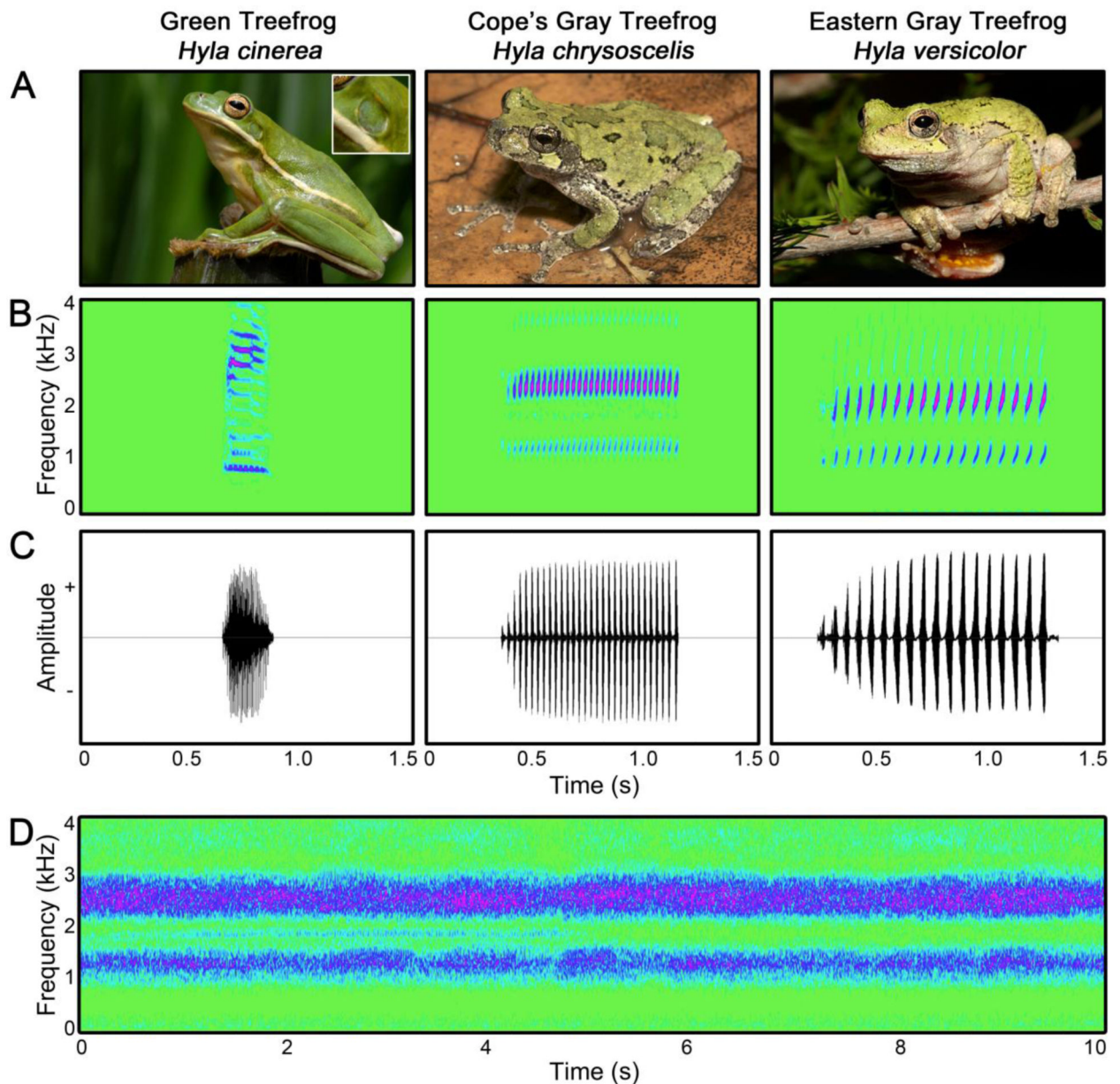
- Simmons, AM.; Moss, CF. Reflex modification: A tool for assessing basic auditory function in anuran amphibians. In: Klump, GM.; Dooling, RJ.; Fay, RR.; Stebbins, WC., editors. *Methods in Comparative Psychoacoustics*. Birkhäuser Verlag; Basel: 1995. p. 197-208.
- Simmons AM, Shen Y, Sanderson MI. Neural and computational basis for periodicity extraction in frog peripheral auditory system. *Audit. Neurosci.* 1996; 2:109–133.
- Simmons, DD.; Meenderink, SWF.; Vassilakis, PN. Anatomy, physiology, and function of the auditory end-organs in the frog inner ear. In: Narins, PA.; Feng, AS.; Fay, RR.; Popper, AN., editors. *Hearing and Sound Communication in Amphibians*. Springer; New York: 2007. p. 184-220.
- Singh NC, Theunissen FE. Modulation spectra of natural sounds and ethological theories of auditory processing. *J. Acoust. Soc. Am.* 2003; 114:3394–3411. [PubMed: 14714819]
- Smith SA, de Oca ANM, Reeder TW, Wiens JJ. A phylogenetic perspective on elevational species richness patterns in Middle American treefrogs: Why so few species in lowland tropical rainforests? *Evolution.* 2007; 61:1188–1207. [PubMed: 17492971]
- Smith SA, Stephens PR, Wiens JJ. Replicate patterns of species richness, historical biogeography, and phylogeny in Holarctic treefrogs. *Evolution.* 2005; 59:2433–2450. [PubMed: 16396184]
- Smotherman MS, Narins PM. Hair cells, hearing and hopping: A field guide to hair cell physiology in the frog. *J. Exp. Bio.* 2000; 203:2237–2246. [PubMed: 10887064]
- Snyder JS, Gregg MK, Weintraub DM, Alain C. Attention, awareness, and the perception of auditory scenes. *Front. Psychol.* 2012;3. [PubMed: 22291676]
- Stosiek C, Garaschuk O, Holthoff K, Konnerth A. *In vivo* two-photon calcium imaging of neuronal networks. *P. Natl. Acad. Sci. USA.* 2003; 100:7319–7324.
- Swanson EM, Tekmen SM, Bee MA. Do female anurans exploit inadvertent social information to locate breeding aggregations? *Can. J. Zool.* 2007; 85:921–932.
- Taylor RC, Klein BA, Stein J, Ryan MJ. Faux frogs: Multimodal signalling and the value of robotics in animal behaviour. *Anim. Behav.* 2008; 76:1089–1097.
- Taylor RC, Ryan MJ. Interactions of multisensory components perceptually rescue túngara frog mating signals. *Science.* 2013; 341:273–274. [PubMed: 23744778]
- van Dijk P, Manley GA. Distortion product otoacoustic emissions in the tree frog *Hyla cinerea*. *Hear. Res.* 2001; 153:14–22. [PubMed: 11223293]
- van Dijk P, Narins PM, Wang JX. Spontaneous otoacoustic emissions in seven frog species. *Hear. Res.* 1996; 101:102–112. [PubMed: 8951437]
- van Noorden, LPAS. Ph.D. thesis, Eindhoven University of Technology. 1975. Temporal coherence in the perception of tone sequences..
- Van Stokkum IHM. Modeling the response of auditory midbrain neurons in the grassfrog to temporally structured monaural stimuli. *Hear. Res.* 1990; 43:231–250. [PubMed: 2312415]
- Van Stokkum IHM, Gielen CCAM. A model for the peripheral auditory nervous system of the grassfrog. *Hear. Res.* 1989; 41:71–85. [PubMed: 2793616]
- Van Stokkum IHM, Melssen WJ. Measuring and modeling the response of auditory midbrain neurons in the grassfrog to temporally structured binaural stimuli. *Hear. Res.* 1991; 52:113–132. [PubMed: 2061201]
- Vélez A, Bee MA. Signal recognition by frogs in the presence of temporally fluctuating chorus-shaped noise. *Behav. Ecol. Sociobiol.* 2010; 64:1695–1709. [PubMed: 21170157]
- Vélez A, Bee MA. Dip listening and the cocktail party problem in grey treefrogs: Signal recognition in temporally fluctuating noise. *Anim. Behav.* 2011; 82:1319–1327. [PubMed: 22389519]
- Vélez A, Bee MA. Signal recognition by green treefrogs (*Hyla cinerea*) and Cope's gray treefrogs (*Hyla chrysoscelis*) in naturally fluctuating noise. *J. Comp. Psychol.* 2013; 127:166–178. [PubMed: 23106802]
- Vélez A, Höbel G, Gordon NM, Bee MA. Dip listening or modulation masking? Call recognition by green treefrogs (*Hyla cinerea*) in temporally fluctuating noise. *J. Comp. Physiol. A.* 2012; 198:891–904.

- Vélez A, Linehan-Skillings BJ, Gu Y, Sun Y, Bee MA. Pulse number-discrimination by Cope's gray treefrog (*Hyla chrysoscelis*) in modulated and unmodulated noise. *J. Acoust. Soc. Am.* 2013; 134:2079–3089.
- Vélez, A.; Schwartz, JJ.; Bee, MA. Anuran acoustic signal perception in noisy environments. In: Brumm, H., editor. *Animal Communication and Noise*. Springer; New York: 2014. in press
- Verhey JL, Pressnitzer D, Winter IM. The psychophysics and physiology of comodulation masking release. *Exp. Brain Res.* 2003; 153:405–417. [PubMed: 13680049]
- Ward JL, Buerkle NP, Bee MA. Spatial release from masking improves sound pattern discrimination along a biologically relevant pulse-rate continuum in gray treefrogs. *Hear. Res.* 2013a; 306:63–75. [PubMed: 24055623]
- Ward JL, Love EK, Vélez A, Buerkle NP, O'Bryan LR, Bee MA. Multitasking males and multiplicative females: Dynamic signalling and receiver preferences in Cope's grey treefrog (*Hyla chrysoscelis*). *Anim. Behav.* 2013b; 86:231–243.
- Warren, RM. *Auditory Perception: A New Analysis and Synthesis*. Cambridge University Press; Cambridge: 1999.
- Watson CS. Some comments on informational masking. *Acta Acust. United Ac.* 2005; 91:502–512.
- Webster, DB.; Fay, RR.; Popper, AN. *The Evolutionary Biology of Hearing*. Springer; New York: 1992. p. 859
- Welch AM, Semlitsch RD, Gerhardt HC. Call duration as an indicator of genetic quality in male gray treefrogs. *Science.* 1998; 280:1928–1930. [PubMed: 9632389]
- Wells, KD. *The Ecology and Behavior of Amphibians*. University of Chicago Press; Chicago: 2007. p. 1148
- Wiens JJ. Global patterns of diversification and species richness in amphibians. *Am. Nat.* 2007; 170:S86–S106. [PubMed: 17874387]
- Wilczynski W, Capranica RR. The auditory system of anuran amphibians. *Prog. Neurobiol.* 1984; 22:1–38. [PubMed: 6374763]
- Wilczynski W, Ryan MJ. The behavioral neuroscience of anuran social signal processing. *Curr. Op. Neurobiol.* 2010; 20:754–763. [PubMed: 20863685]
- Wotton JM, Ferragamo MJ. A model of anuran auditory periphery reveals frequency-dependent adaptation to be a contributing mechanism for two-tone suppression and amplitude modulation coding. *Hear. Res.* 2011; 280:109–121. [PubMed: 21565263]
- Zakon, HH.; Wilczynski, W. The physiology of the anuran eighth nerve. In: Fritzsche, B.; Wolkowiak, W.; Ryan, MJ.; Wilczynski, W.; Hetherington, T., editors. *The Evolution of the Amphibian Auditory System*. Wiley; New York: 1988. p. 125-155.



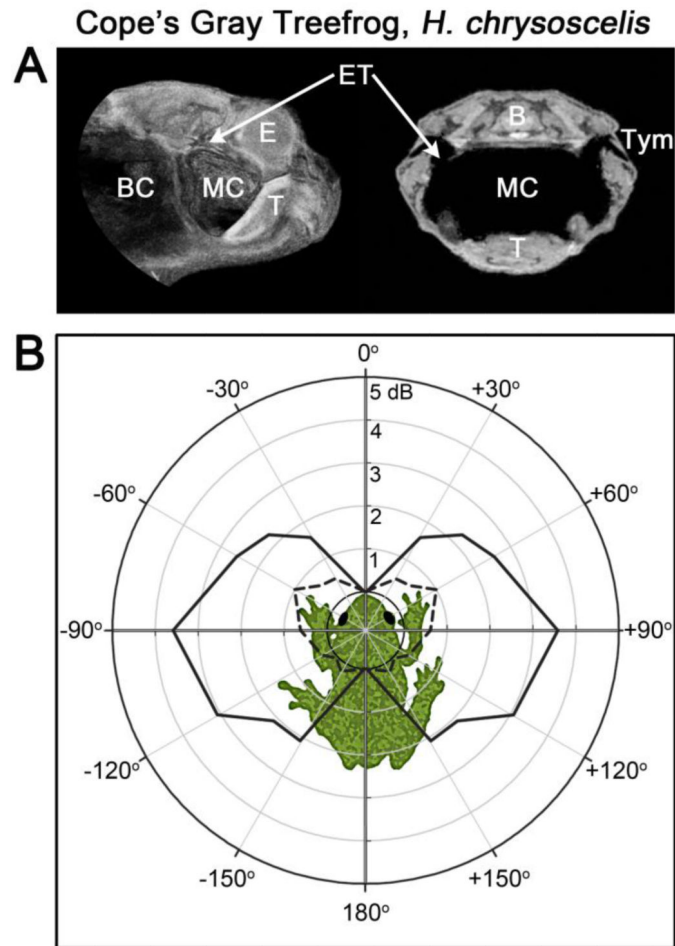
### Highlights

- Treefrogs are ethological specialists at communicating in noisy social groups.
- The treefrog's cocktail-party-like problem is reviewed in an evolutionary context.
- Previous and ongoing research on auditory scene analysis in treefrogs is reviewed.
- Auditory grouping, spatial release from masking and dip listening are discussed.
- A roadmap for future work on scene analysis with this animal model is outlined.

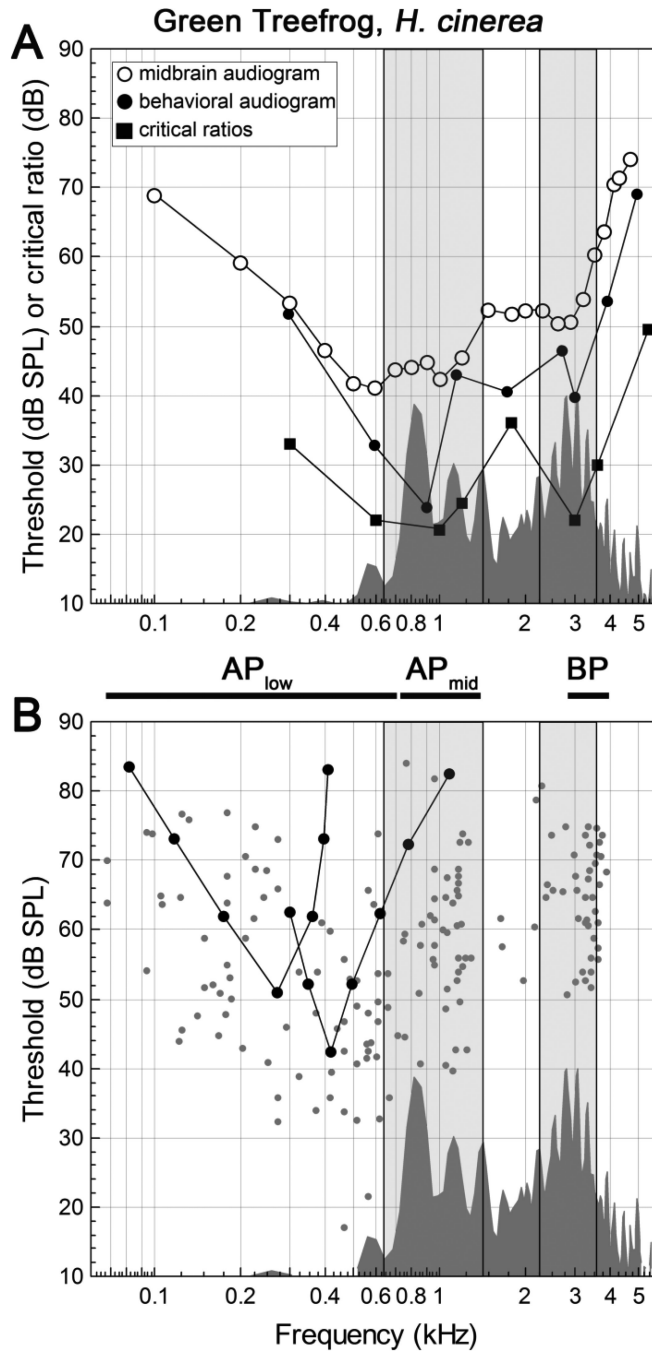
**Fig. 1.**

Three well-studied treefrog species in the genus *Hyla* and depictions of their advertisement calls and a treefrog chorus. **A** Photographs of adults of the green treefrog (*H. cinerea*), Cope's gray treefrog (*H. chrysoscelis*), and the eastern gray treefrog (*H. versicolor*). *Inset*: magnified view of the tympanum of the green treefrog. **B** Spectrograms (1024 pt. FFT) of advertisement calls showing frequency as a function of time, with relative amplitude depicted by color intensity. Note that each call has a bimodal frequency spectrum. In green treefrogs, advertisement calls have acoustic energy in the region of 0.9 kHz and 2.5–3.5 kHz; in the two gray treefrogs, advertisement calls have spectral energy in the range of 1.1–1.4 kHz and 2.2–2.8 kHz. In all three species, the low range of frequencies is primarily encoded by the amphibian papilla, whereas the upper range of frequencies is primarily

encoded by the basilar papilla. **C** Waveforms depicting the amplitude envelope of single calls. Note that the green treefrog advertisement call consists of a single note (~150–250 ms in duration), whereas the calls of the two gray treefrogs are longer and pulsatile, with pulse rates of about 35–50 pulses/s in Cope's gray treefrog and 18–24 pulses/s in eastern gray treefrogs (at 20° C). During a typical night of chorusing, males produce calls like those depicted here for long periods (e.g., for several hours) at high repetition rates. In green treefrogs, males commonly produce the single-note call at rates of about one call every 1–2 s, while the males of both gray treefrogs species produce their pulsatile calls at rates of about 1 call every 5–15 s. **D** Spectrogram (1024 pt. FFT) of a 10-s recording of a dense chorus of Cope's gray treefrogs illustrating two continuous bands of noise characteristic of choruses of conspecifics. This chorus was recorded in a small pond in central Minnesota. The exact number of males calling in this chorus is unknown, but probably ranges between 100 to 300 males. As illustrated here, there is considerable overlap between the calls of males in the chorus, which makes a female's task of choosing her mate a difficult cocktail-party-like problem. Photo credits: *Hyla cinerea* © 2007 John White, *Hyla chrysoscelis* and *Hyla versicolor* © 2010 Todd Pierson / [www.discoverlife.org](http://www.discoverlife.org).



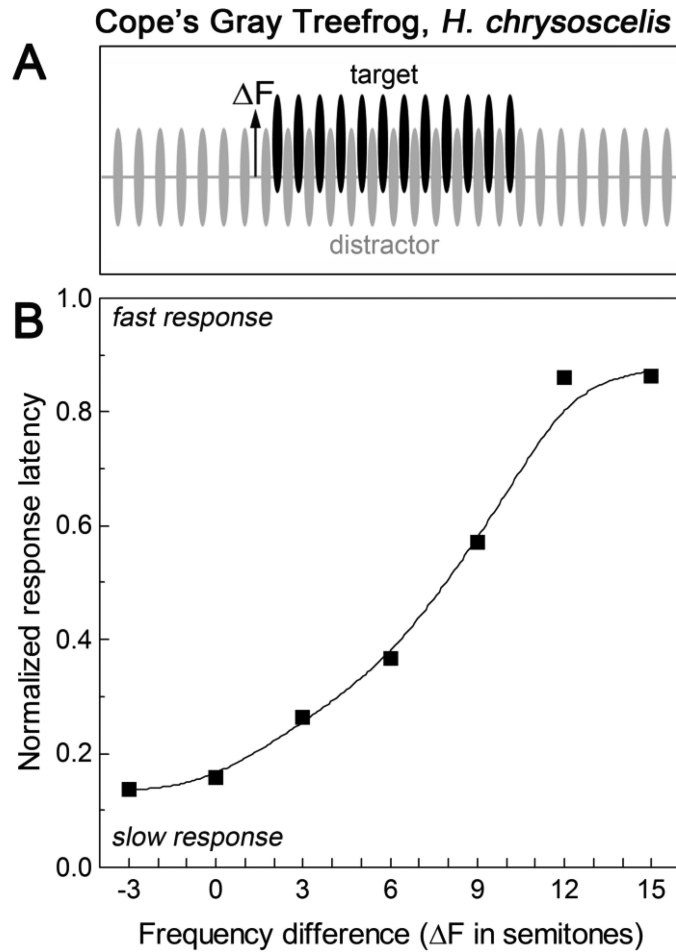
**Fig. 2.** Spatial hearing in Cope's gray treefrog, *Hyla chrysoscelis*. **A** Magnetic resonance imaging (MRI) scans of a gray treefrog showing the air passageways that allow for coupling of the left and right tympana (Tym) through the air-filled middle ear cavities, the wide Eustachian tubes (ET), and the mouth cavity (MC) (E = eye, T = tongue, BC = body cavity, B = brain). The left image shows a lateral view through a 3-dimensional image stack and the right image shows a single coronal slice passing through the middle ears, Eustachian tubes, and mouth. Note the close proximity between the mouth cavity (MC) and body cavity (BC). MRI scans were made with a 9.4 Tesla magnet with 31-cm bore. **B** Difference (in dB) between the two ears in the amplitude of a call (dashed lines) and the tympanum's response to the same call (solid lines) as a function of sound incident angle in the azimuthal plane (data redrawn from Caldwell et al., submitted). The inter-aural difference in the amplitude of the call was measured using a probe microphone placed adjacent to the tympanum. The inter-aural difference in the tympanum's response was measured using laser Doppler vibrometry. Both types of measurements were made on the same, single side of the frog as the sound presentation was moved through 360° over 16 measurement angles. The data have been folded over the midline to illustrate the inherent directionality that results from having inputs involving internally coupled tympana and extratympanic pathways.



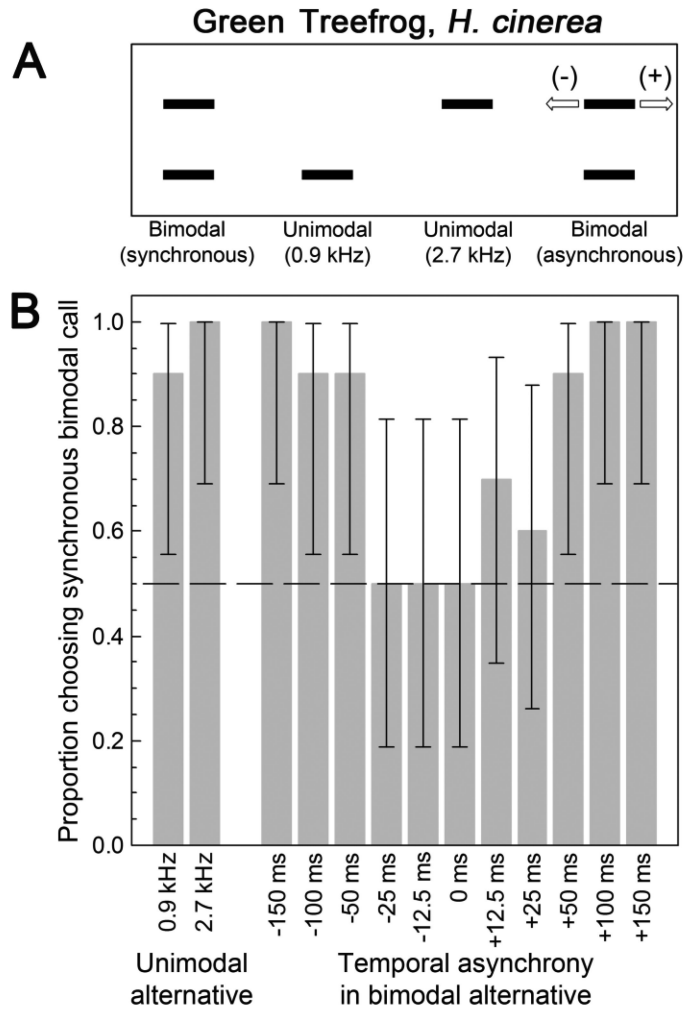
**Fig. 3.** Frequency tuning in the green treefrog, *Hyla cinerea*. **A** Shown here are a midbrain audiogram (open circles), a behavioral audiogram (solid circles), and critical ratios as a function of frequency (solid squares). The midbrain audiogram was computed by averaging audiograms derived from multiunit responses recorded in the torus semicircularis (inferior colliculus) reported in Figure 5 of Lombard and Straughan (1974), Figure 1A in Miranda and Wilczynski (2009b), and Figures 4A and 4C in Penna et al. (1992). Linear interpolation between adjacent frequencies was used for a few values to compensate for different stimulus

frequencies across studies. The behavioral audiogram represents an average of audiograms from two individuals determined using reflex modification, redrawn from Figure 6 in Megela-Simmons et al. (1985). Critical ratios were determined using reflex modification and tone presentations in noise (35 dB spectrum level) and are redrawn from Figure 6 in Moss and Simmons (1986). Between plots **A** and **B** are solid lines showing the approximate distributions of the low-frequency ( $AP_{low}$ ) and mid-frequency ( $AP_{mid}$ ) tuning of the amphibian papilla and the tuning of the basilar papilla (BP) based on single-unit recordings of auditory nerve fibers (after Ehret and Capranica, 1980). **B** Distribution of best excitatory frequencies and corresponding thresholds (gray dots) of single auditory nerve fibers, redrawn from Figure 1A in Ehret and Capranica (1980), and excitatory tuning curves (solid squares) from two auditory nerve fibers in the  $AP_{low}$  range showing typical V-shaped tuning, redrawn from (left) Figure 6 in Capranica and Moffat (1983) and (right) Figure 1 in Ehret et al. (1983). The *insets* in **A** and **B** show the power spectrum of the *H. cinerea* advertisement call depicted in Figure 1. The two shaded areas in each plot depict the range of frequencies across individuals reported for the low-frequency and high-frequency spectral peaks of the call, as reported in Gerhardt (2001). Small differences on the order of 100 to 200 Hz between peaks in the call and peaks in neural and behavioral sensitivity and selectivity almost certainly reflect the influences of differences in temperature at the times recordings were made and geographic variation across studies.

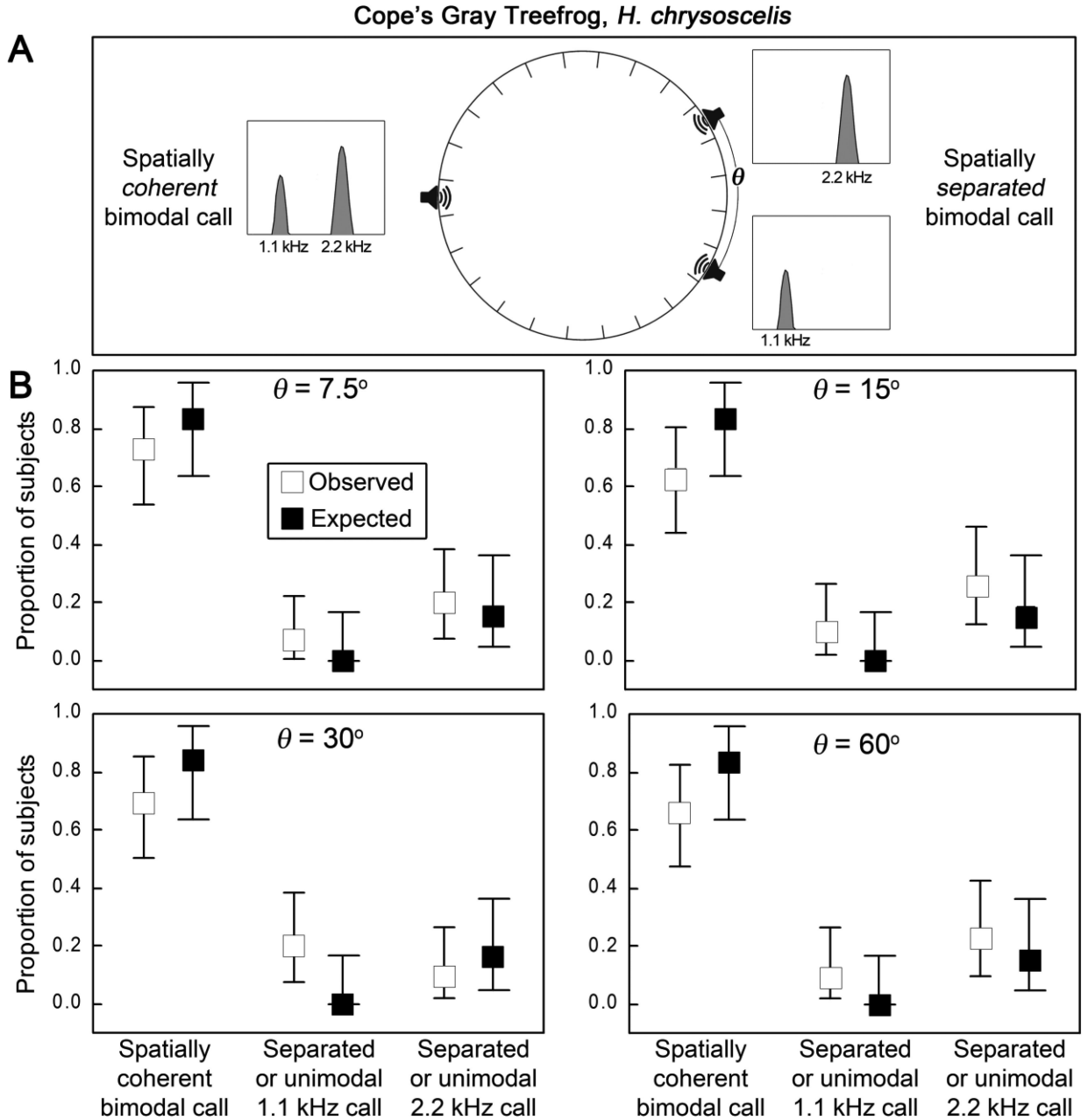




**Fig. 4.** Spectral proximity ( $\Delta F$ ) as a cue for sequential integration in Cope's gray treefrog, *Hyla chrysoscelis*. **A** Schematic diagram illustrating the temporal relationships of pulses in the temporally discrete target signal (45.5 pulses/s) and a continuous, pulsatile distractor (45.5 pulses/s). The pulses of the two sounds were temporally interleaved to create a composite pulse rate of 91 pulses/s when the target played during the distractor. The frequency of the target was fixed (1.3 kHz or 2.6 kHz) and the frequency of the distractor was varied across trials to achieve the nominal frequency separation ( $\Delta F$ ), illustrated here by shifting the target upward from the distractor. **B** Points depict normalized response latencies (1 = fast response, 0 = slow/no response) as a function of frequency separation ( $\Delta F$  in semitones) between the target and distractor. A value of 1.0 corresponds to the latency to approach an attractive synthetic call presented by itself. Data redrawn from Nityananda and Bee (2011).



**Fig. 5.** Common onsets and offsets as a possible cue for simultaneous integration in green treefrogs, *Hyla cinerea*. **A** Schematic spectrograms showing a synthetic bimodal call with two synchronous spectral peaks at 0.9 kHz and 2.7 kHz, two unimodal calls with just a 0.9 kHz or the 2.7 kHz spectral peak, and a bimodal call in which temporal asynchrony was introduced by advancing (-) or delaying (+) the onset and offset of the 2.7 kHz spectral peak relative to those of the 0.9 kHz spectral peak. The call duration was 150 ms and the degree of temporal asynchrony across stimuli ranged from -150 ms to +150 ms. **B** Preliminary results from two-stimulus choice tests showing the mean proportion ( $\pm$  exact binomial confidence intervals) of subjects ( $N = 10$ ) that chose a synchronous bimodal call when it was paired against the alternative indicated along the x-axis. The horizontal dashed line indicates the expected proportion based on chance in a two-stimulus choice test.



**Fig. 6.** Spatial proximity as a cue for simultaneous integration in Cope's gray treefrog, *Hyla chrysoscelis*. **A** Schematic diagram of the test arena (2 m diameter) showing speaker arrangements for the critical two-stimulus choice tests in which a spatially coherent bimodal call (left) alternated in time with a spatially separated bimodal call (right), in which the two spectral peaks were presented simultaneously from speakers separated by an angle ( $\theta$ ) of 7.5°, 15°, 30°, or 60° in separate tests. **B** Proportion of subjects ( $\pm$  95% exact binomial confidence intervals) that chose the spatially coherent bimodal call over the separated bimodal alternative (open squares). These data are depicted in comparison with expected proportions (solid squares) based on control tests in which the spatially coherent bimodal

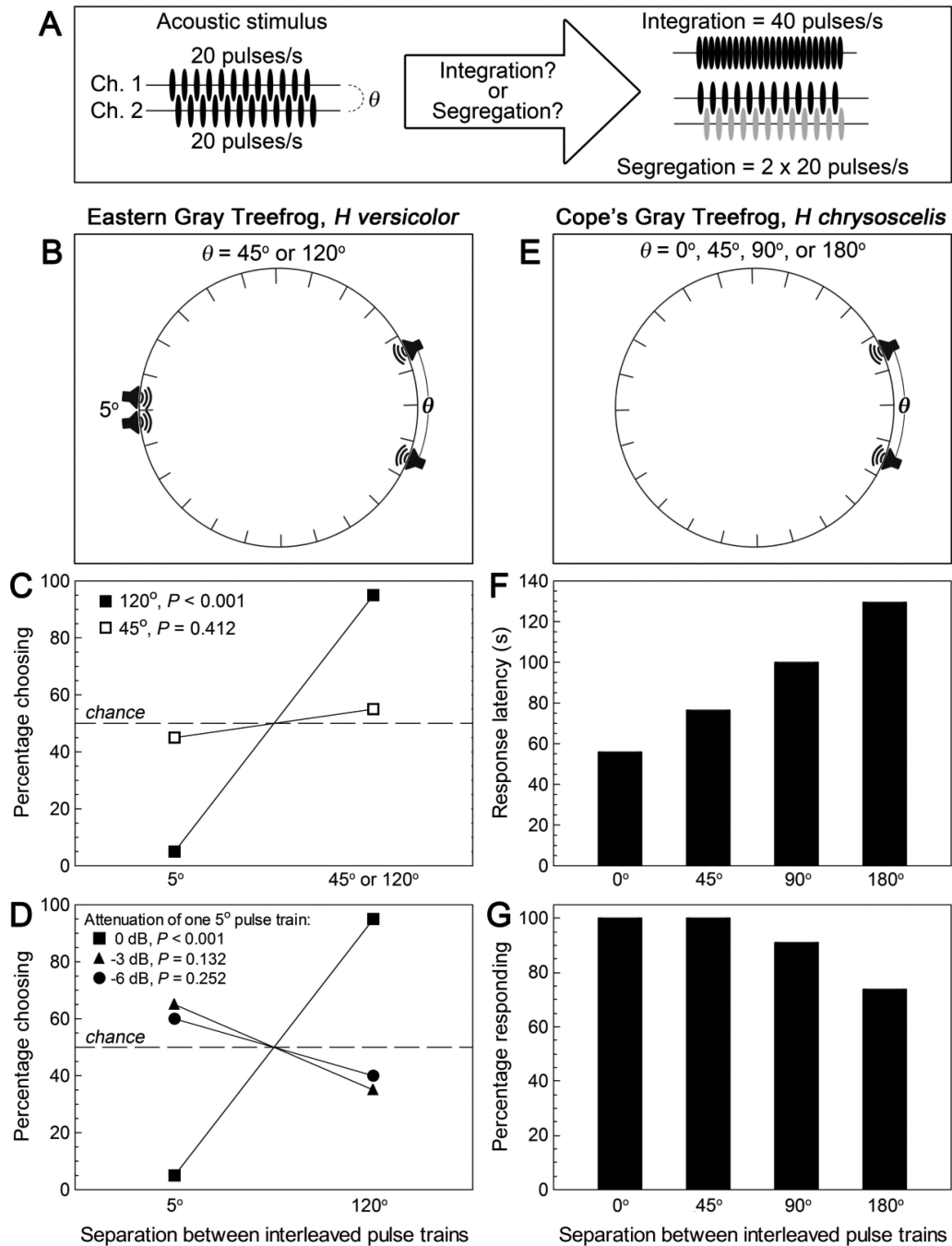
call was paired against a unimodal alternative from the opposite side of the arena. Figure adapted from Bee, M.A. 2010. Spectral preferences and the role of spatial coherence in simultaneous integration in gray treefrogs (*Hyla chrysoscelis*). *Journal of Comparative Psychology* 124, 412-424, with permission from the American Psychological Association.

Author Manuscript

Author Manuscript

Author Manuscript

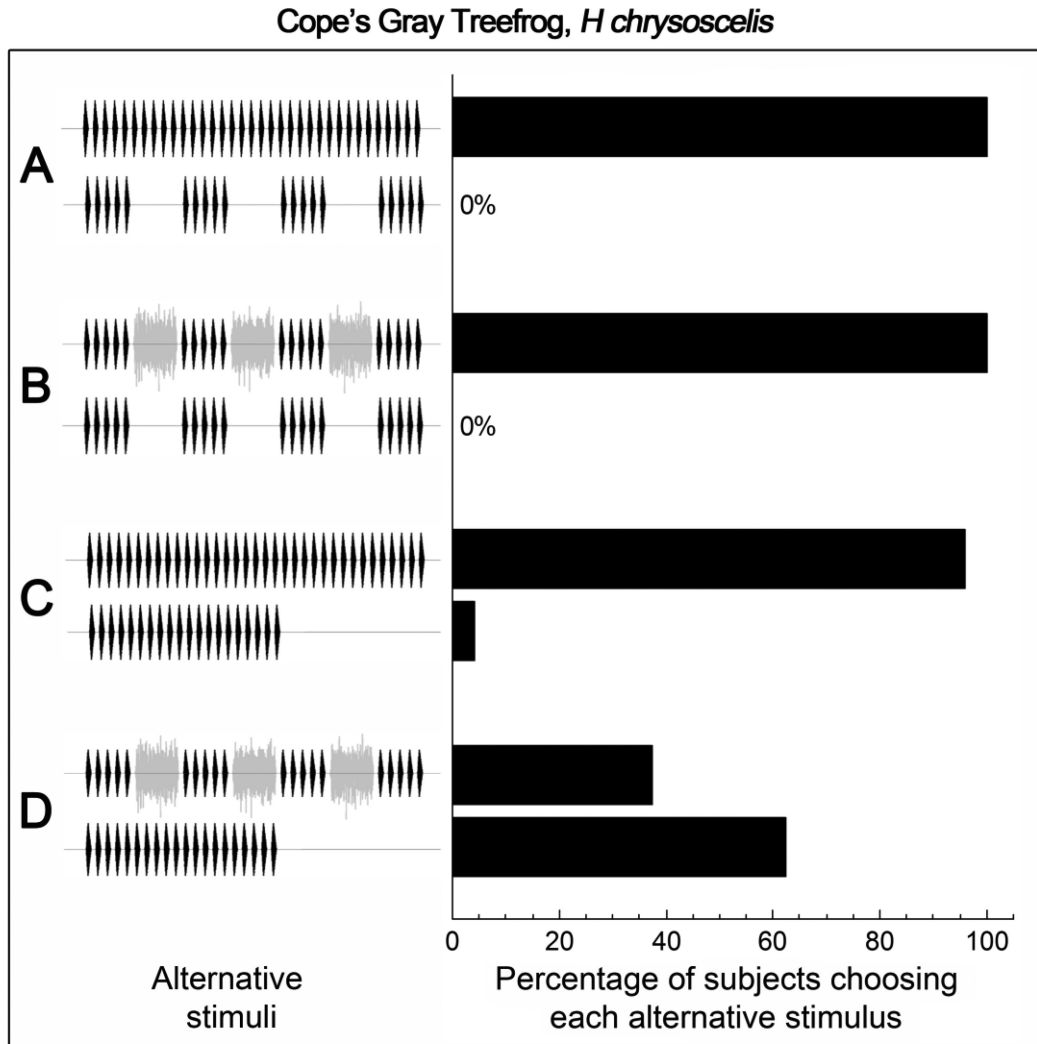
Author Manuscript



**Fig. 7.** Spatial proximity as a cue for sequential integration in the eastern gray treefrog, *Hyla versicolor* and Cope's gray treefrog, *Hyla chrysoscelis*. **A** Schematic diagram illustrating how interleaved pulses were used as stimuli. Each separate pulse train had a pulse rate in the range of *Hyla versicolor* (e.g., 20 pulses/s) and was presented from one of two spatially separated speakers on different stereo channels to create a single call with temporally interleaved pulses having a composite pulse rate in the range of *Hyla chrysoscelis* (e.g., 40 pulses/s). Hence, sequential integration of the two pulse trains results in a *Hylachrysoscelis*-

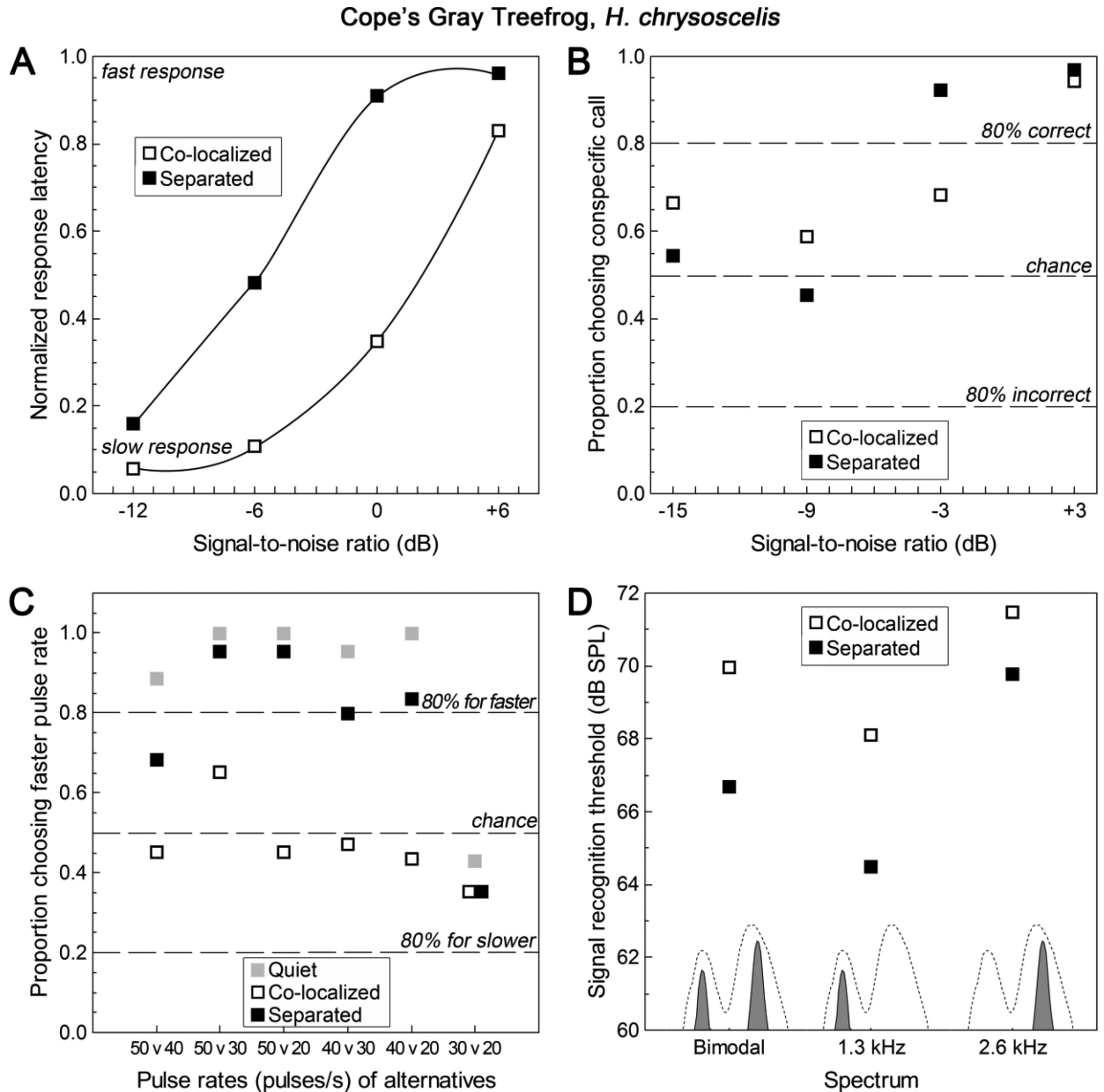
like call (e.g., 40 pulses/s), whereas perceptual segregation results in two *Hyla-versicolor*-like calls (e.g.,  $2 \times 20$  pulses/s). **B-D** Tests of sequential integration in two-stimulus choice tests with *Hyla versicolor* by Schwartz and Gerhardt (1995). **B** Schematic diagram of the test arena (2 m diameter) showing speaker arrangements, in which one alternative was always presented from two speakers separated by  $5^\circ$  and the other was presented from two speakers separated by  $45^\circ$  or  $120^\circ$ . **C** Preference functions for two-stimulus choice tests in which the alternative to  $5^\circ$  separation was separated by either  $45^\circ$  or  $120^\circ$ . **D** Preference functions for two-stimulus choice tests in which one of the two pulse trains presented at  $5^\circ$  separation was attenuated by 0 dB,  $-3$  dB, or  $-6$  dB relative to the other pulse train at  $5^\circ$  separation and the two pulse trains at  $120^\circ$  separation. **E-G** Tests of sequential integration in single-stimulus tests with *Hyla chrysoscelis* by Bee and Riemersma (2008). **E** Schematic diagram of the test arena (2 m diameter) showing speaker arrangements across conditions, in which interleaved pulse trains were presented from a single speaker ( $0^\circ$ ) or two speakers separated by  $45^\circ$ ,  $90^\circ$ , or  $180^\circ$ . **F** Response latency as a function of spatial separation between the two interleaved pulse trains. **G** Percentage of subjects eventually responding as a function of spatial separation between the two interleaved pulse trains. Data redrawn from Schwartz and Gerhardt (1995) and Bee and Riemersma (2008).





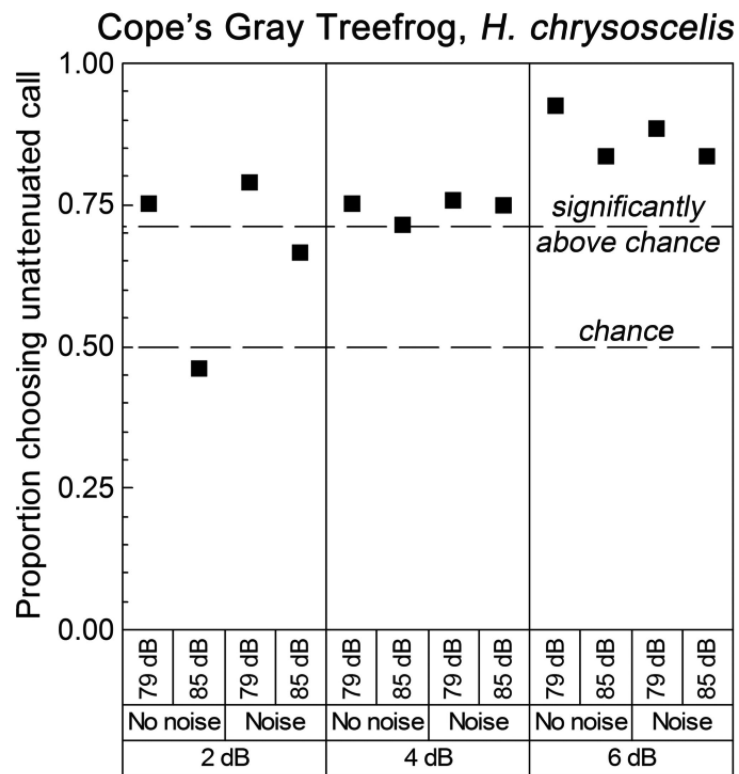
**Fig. 8.**

Testing the continuity illusion in Cope's gray treefrog, *Hyla chrysoscelis*. **A** In two-stimulus choice tests, females discriminated unanimously against a gap call with 20 pulses in favor of a continuous 35-pulse call of equivalent duration. Similar unanimous discrimination against the gap call was also observed when it also had 35 pulses (Seeba et al., 2010). **B** Females unanimously chose a call having gaps filled with noise over a gap call. **C** Given a choice between two continuous calls, females preferred a longer with 35 pulses over a shorter call with 20 pulses. **D** In the critical test of auditory induction, females failed to show significant preferences for a gap-filled call with 20 pulses and duration equivalent to a 35-pulse call over a continuous 20-pulse call. See Seeba et al. (2010) for additional control tests and further discussion. Data and stimuli redrawn from Seeba et al. (2010).

**Fig. 9.**

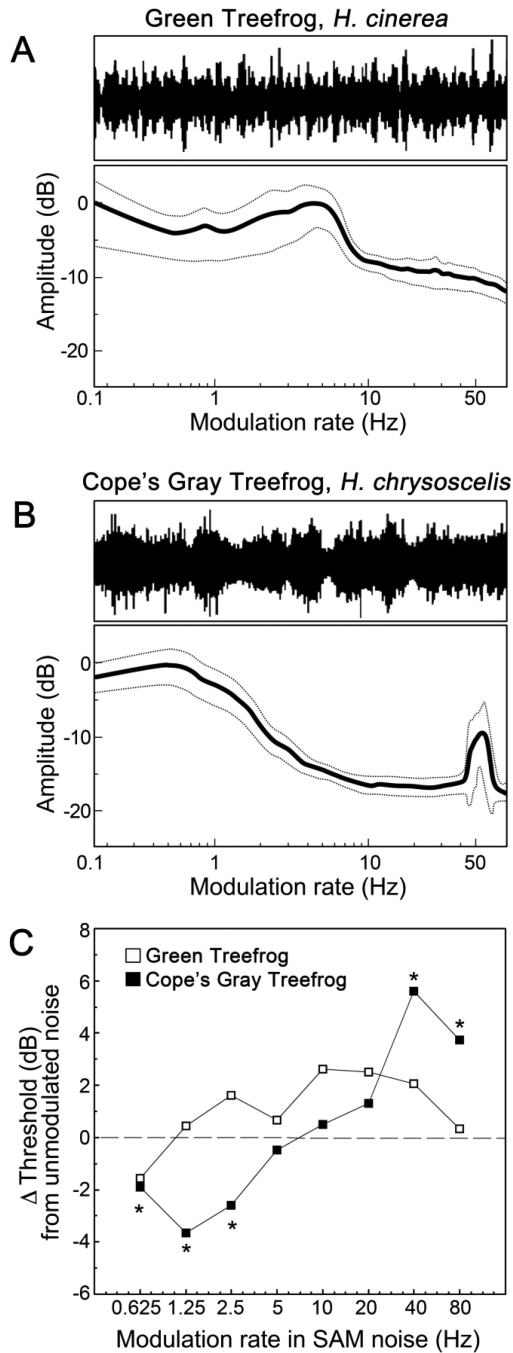
Spatial release from masking in Cope's gray treefrog, *Hyla chrysoscelis*. **A** Normalized reaction times (1 = fast response, 0 = slow/no response) as a function of the SNR in a single-stimulus test in which a source of advertisement calls was co-localized (adjacent) with a source of chorus-shaped noise (open squares) or separated by 90° (solid squares) (Bee, 2007). **B** Proportions of females in a two-stimulus choice test choosing a conspecific call over a heterospecific (*H. versicolor*) call as a function of the SNR in co-localized (speakers adjacent) (open squares) and separated conditions (speakers 90° apart) (solid squares) (Bee, 2008a). **C** Proportions of females in a two-stimulus choice test choosing the alternative with a faster pulse rate in which the two alternatives in a test differed along a continuum of pulse rates between conspecific (50 pulses/s) and heterospecific (20 pulses/s). Data are shown

separately for tests conducted in quiet (gray squares), and in the presence of co-localized (open squares) and 90° separated (solid squares) chorus-shaped noise (Ward et al., 2013a). **D** Signal-recognition thresholds determined in co-localized (open squares) and 90° separated (solid squares) chorus shaped noise as a function of the spectral content of the signal (Nityananda and Bee, 2012). The insets show the spectra of the chorus-shaped noise (dashed line) and each target signal (gray area) with spectral peaks at 1.3 kHz, 2.6 kHz, or both (bimodal). The y-axis for the spectra is approximately 30 dB. Data redrawn from Bee (2007, 2008a), Ward et al. (2013a), and Nityananda & Bee (2012).



**Fig. 10.**

Level discrimination in Cope's gray treefrog, *Hyla chrysoscelis*. Points depict the proportions of subjects ( $N = 24$ ) in two-stimulus choice tests that chose an unattenuated call at the nominal signal level (79 dB or 85 dB SPL) in the presence or absence of chorus-shaped noise (73 dB SPL) when the alternative call was attenuated by 2 dB, 4 dB, or 6 dB. Data redrawn from Bee et al. (2012).



**Fig. 11.** Tests of dip listening using sinusoidally amplitude modulated (SAM) chorus-shaped noise in green treefrogs, *Hyla cinerea*, and Cope's gray treefrog, *Hyla chrysoscelis*. **A** and **B** depict waveforms of 10-s segments of natural choruses (top) and the mean (solid line,  $\pm 1$  standard deviation, dashed lines) modulation spectra (bottom) of the amplitude envelopes of chorus recordings ( $N = 25$  for each species) for green treefrogs (**A**) and Cope's gray treefrogs (**B**). Modulation spectra are based on computing a fast Fourier transformation of the Hilbert envelope of the chorus noise. See additional details in Vélez and Bee (2013). **C** Masked

signal-recognition thresholds in SAM chorus-shaped noise, computed as the threshold difference from an unmodulated control noise and shown as a function of the rate of sinusoidal modulation in the masker. Threshold differences significantly lower or higher than 0 dB are depicted with asterisks. Waveforms and modulation spectra in **A** and **B** adapted from Vélez, A., Bee, M.A. 2013. Signal recognition by green treefrogs (*Hyla cinerea*) and Cope's gray treefrogs (*Hyla chrysoscelis*) in naturally fluctuating noise. *Journal of Comparative Psychology* 127, 166-178, with permission from the American Psychological Association. Data in **C** are redrawn from Vélez et al. (2012) and Vélez and Bee (2011).



**Table 1**

Estimated divergence times between humans (*Homo sapiens*) and several animal models used in modern hearing research. Data represent median estimates of divergence times (in millions of years before present) as reported by Timetree.org on June 19, 2013 (Hedges et al., 2006).

Class	Order	Common name (Genus)	Divergence time	
Mammalia	Primates	rhesus monkeys ( <i>Macaca</i> )	26.8	
		Rodentia	mouse ( <i>Mus</i> ) rat ( <i>Rattus</i> ) Guinea pig ( <i>Cavia</i> ) Mongolian gerbil ( <i>Meriones</i> ) chinchilla ( <i>Chinchilla</i> )	94.5
	Lagomorpha	rabbit ( <i>Oryctolagus</i> )	94.5	
	Carnivora	cat ( <i>Felis</i> )	96.2	
		ferret ( <i>Mustella</i> )		
	Chiroptera	bats (e.g., <i>Eptesicus</i> )	96.2	
	Aves	Strigiformes	barn owl ( <i>Tyto</i> )	322.4
			European starling ( <i>Sturnus</i> ) zebra finch ( <i>Taeniopygia</i> )	322.4
		Psittaciformes	budgerigar ( <i>Melopsittacus</i> )	322.4
	Reptilia	Squamata	tokay gecko ( <i>Gekko</i> )	322.4
Amphibia	Anura	African clawed frog ( <i>Xenopus</i> )	359.1	
		northern leopard frog ( <i>Rana</i> ) green treefrog ( <i>Hyla</i> )		
		Actinopterygii	Cypriniformes	goldfish ( <i>Carassius</i> ) zebrafish ( <i>Danio</i> )
	Batrachoidiformes	toadfish ( <i>Opsanus</i> )	436.8	
		midshipman ( <i>Porichthys</i> )	436.8	
Insecta	Diptera	fruit fly ( <i>Drosophila</i> )	725.5	
		cricket parasitoid ( <i>Ormia</i> )		