Sound Source Perception in Anuran Amphibians

Mark A Bee

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

Abstract

Sound source perception refers to the auditory system's ability to parse incoming sensory information into coherent representations of distinct sound sources in the environment. Such abilities are no doubt key to successful communication in many taxa, but we know little about their function in animal communication systems. For anuran amphibians (frogs and toads), social and reproductive behaviors depend on a listener's ability to hear and identify sound signals amid high levels of background noise in acoustically cluttered environments. Recent neuroethological studies are revealing how frogs parse these complex acoustic scenes to identify individual calls in noisy breeding choruses. Current evidence highlights some interesting similarities and differences in how the auditory systems of frogs and other vertebrates (most notably birds and mammals) perform auditory scene analysis.

Introduction

Humans and other animals often communicate acoustically in large social aggregations. Familiar examples might include a cocktail party, a songbird dawn chorus, or a group of singing insects. An important problem in the study of animal communication and sensory neurobiology concerns how a receiver's sense of hearing enables them to perceive acoustic signals as distinct sounds in these noisy and acoustically cluttered environments [1, 2]. Sound pressure waves produced by multiple active sources add together to form a single, composite sound pressure waveform that impinges on the receiver's ears (Figure 1). Listeners must parse this complex sensory input to make sense of the surrounding acoustic scene, and this ability involves overcoming two related challenges. One of these challenges is auditory masking. Signalers cannot be perceived as distinct sound sources unless receivers can segregate signals from the ambient background noise. The second challenge involves perceptually binding the spectral and temporal sound elements composing an acoustic signal to create a coherent auditory percept of the signal that can be assigned to the correct source. For humans, these two challenges form the basis of the so-called "cocktail party problem," which refers to the difficulty we experience listening to speech in noisy social settings (Figure 1) [3].

Compared with our well-developed understanding of sound source perception in humans [4–8], we know less about how non-human animals perform similar tasks in the general context of hearing and in the more specific context of acoustic communication [1, 2, 9]. This

Correspondence: Mark A. Bee, Department of Ecology, Evolution and Behavior, 1987 Upper Buford Circle, University of Minnesota, St. Paul, MN 55108, U.S.A., Ph: 612-624-6749, Fx: 612-624-6777, mbee@umn.edu.

taxonomic disparity in knowledge about sound source perception represents an exciting research opportunity for neuroethologists. The sense of hearing had multiple evolutionary origins [10], and key features of auditory processing within some lineages arose multiple times (e.g., tympanic hearing in terrestrial vertebrates) [11, 12]. From an evolutionary perspective, comparative neuroethological studies of diverse animal groups are necessary to understand how evolution might have solved common problems in sound source perception in different lineages [12]. Here, I review studies of sound source perception in anuran amphibians, a group of vertebrates for which acoustic communication is the most important form of socio-sexual communication.

The frog's cocktail party problem

Anurans are notable for the loud vocalizations males produce to attract females and to defend calling sites against rival males [13, 14]. Frog vocalizations commonly reach peak sound pressure levels (SPLs) of 90 dB to 110 dB (re 20 μPa; measured at 1 m) [15]. In many species, communication takes place in large breeding choruses comprising hundreds of males, usually of multiple species, gathered at a suitable breeding site (e.g., a pond). Ambient chorus noise is intense [16], and has been reported audible to humans up to distances of 2 km from a breeding pond [17]. High noise levels and temporally overlapping calls within a chorus interfere with the ability of receivers to detect, recognize, and discriminate among male vocalizations [18]. Nevertheless, female frogs choose mates of their own species based on the spectrotemporal properties of his calls and can exercise more selective choices of particular conspecific males producing calls with certain properties [13, 14]. Male frogs can assess the proximity, size, fighting ability, and even individual identity of other calling males in a chorus environment [13, 14]. These behaviors indicate frogs perceive calling males as distinct sound sources in noisy chorus environments. Research is beginning to uncover how frogs hear individual sound sources under these real-world listening conditions [18, 19].

Segregating signals from masking noise

In humans, exploiting spatial separation between speech signals and noise represents one important way we ameliorate auditory masking under cocktail-party-like listening conditions [20]. Frogs appear to do so as well. Schwartz and Gerhardt [21] showed that female green treefrogs, *Hyla cinerea*, responded to male calls at signal-to-noise ratios (SNRs) that were about 3 dB lower when there was spatial separation between signals and noise, although spatial separation did not improve discrimination between advertisement calls and aggressive calls. Females of Cope's gray treefrog, *Hyla chrysoscelis*, experienced about 4 dB to 6 dB of masking release when signals and "chorus-shaped noise" were spatially separated (Figure 2a) [22, 23]. Spatial release from masking also improved the ability of females of this species to discriminate between conspecific calls and the acoustically similar calls of a closely related congener [24].

At a neurophysiological level, lower neural signal detection thresholds have been reported to occur under spatially separated conditions in studies of spatial release from masking in northern leopard frogs, *Lithobates* (formerly *Rana*) *pipiens* [25–27]. The magnitude of spatial unmasking observed in neural responses increased from an average maximum of about 3 dB in auditory nerve fibers to about 9 dB in the torus semicircularis, the midbrain region homologous to the inferior colliculus [26]. Using iontophoretic applications of bicuculline, a $GABA_A$ receptor antagonist, Lin and Feng [27] showed that GABAergic interactions were, in part, responsible for the more robust spatial release observed in some midbrain neurons when compared with the periphery. Thus, central processing mechanisms, such as binaural inhibition, contribute to spatial release from masking in frogs.

Humans also exploit temporal fluctuations in the level of speech-like masking noise to hear speech and other sounds at lower SNRs [28]. This has been attributed to our ability to "listen in the dips" of fluctuating noise to catch acoustic glimpses of sounds during brief periods of low noise amplitude. Some frogs are also capable of listening in the dips. Using sinusoidally amplitude modulated (SAM) noise with the spectrum of natural breeding choruses, we recently found that female Cope's gray treefrogs, *H. chrysoscelis*, experienced about 2 dB to 4 dB of masking release when the masker fluctuated at slow rates (e.g., 0.625 Hz to 2.5 Hz; Figure 2b) [29]. These rates were precisely the same slow rates of fluctuations present in the modulation spectra of natural chorus sounds [29, 30]. Interestingly, masking release was only observed when the "dips" in the masker were long enough for about nine or more consecutive pulses to "fit" in a dip. A follow-up experiment indicated that a threshold number of 6 to 9 pulses was required to elicit phonotaxis in quiet conditions. Together, these findings suggest that females were better able to catch meaningful acoustic glimpses of calls having a sufficient number of pulses in maskers fluctuating at slow rates. It seems likely, though presently this is pure speculation, that dip listening for pulsatile calls involves neurons in the frog midbrain that "count" pulses based on integrating temporal information over a number of correct interpulse intervals [31].

Psychophysical estimates of masking release in frogs as a result of spatial unmasking or dip listening is typically on the order of 2 dB to 6 dB (Figure 2). While somewhat smaller than that found in humans, these magnitudes of masking release are large enough to be biologically important for frogs. Female frogs exhibit robust discrimination of differences in level as small as 2 dB to 4 dB both in quiet and in the presence of chorus noise [32] and they may also *reverse* their directional preferences for calls with certain acoustic properties when the normally preferred signal is attenuated by similar magnitudes [13]. An important goal for future research will be to investigate the extent of similarities and differences between frogs and other animals in the underlying neural mechanisms responsible for exploiting spatial and temporal cues in segregating communication signals from noise.

Auditory grouping

Frog calls and other animal acoustic signals commonly consist of spectrally rich sounds arranged in sequences of temporally discrete elements, such as pulses or notes. Auditory grouping, therefore, requires both "simultaneous integration" (i.e., grouping concurrent harmonics or formants) and "sequential integration" (i.e., grouping temporally separated elements through time) [4]. Rigorous psychoacoustic studies of human subjects suggest the general rule that sound elements sharing features in common tend to be grouped together into coherent "auditory objects" or "auditory streams" [4–8]. Among the most potent auditory grouping cues are common onsets/offsets, common fundamental frequency F_0 (i.e., "harmonicity" or spectral regularity), and common fluctuations in amplitude. Common spatial origin is another potential grouping cue, though sounds coming from different directions can sometimes be grouped together based on commonalities in their acoustic properties [33].

Simultaneous integration

Early studies of simultaneous integration in frogs focused on common F_0 as a potential auditory grouping cue and the results are mixed. North American bullfrogs, *Lithobates catesbeianus* (formerly *Rana catesbeiana*), produce a low-frequency, moan-like call consisting of a series of 10–20 prominent harmonics of a F_0 of about 90–120 Hz. A field playback study by Simmons and Bean [34] showed that territorial males perceive calls with a mistuned harmonic as perceptually distinct from purely harmonic complexes. This finding indicates that common F_0 could function as a cue for simultaneous integration in this species. In a reflex modification study of green treefrogs, *H. cinerea*, Simmons et al. [35]

found an 8 dB to 10 dB difference in the masked thresholds for detecting harmonic versus inharmonic complex tones, indicating sensitivity to harmonicity in this species as well. Subsequent studies, however, failed to find strong evidence that either male or female green treefrogs discriminated between harmonic and inharmonic calls [36, 37], though females of the related barking treefrog, *H. gratiosa,* did discriminate among harmonic and inharmonic calls in similar experiments [38].

One recent study investigated the effects of common spatial origin on simultaneous integration in frogs. Males of Cope's gray treefrog, *H. chrysoscelis,* produce a pulsatile call, with each pulse comprising two harmonically-related spectral elements of about 1.25 kHz and 2.5 kHz. In laboratory tests of phonotaxis behavior, females were given a choice between spatially coherent signals (both harmonics from one speaker) and spatially incoherent alternatives (each harmonic from a different speaker). Females exhibited robust discrimination against calls with spatially incoherent harmonics separated by as little as 7.5°, a finding consistent with the hypothesis that common spatial origin promotes auditory grouping of the two harmonics [39].

Sequential integration

Several studies of auditory grouping in frogs have focused on common spatial origin as a potential grouping cue for sequential integration of temporally separated call elements (Figure 3) [40–45]. These studies reveal that frogs are fairly permissive of spatial incoherence and readily group sequential sounds arising from different locations. In the Australian quacking frog, *Crinia georgiana,* for example, males produce a multi-note vocalization composed of a sequence of "quacks." During interactions in the chorus, males attempt to competitively match the number of quacks in a neighbor's calls. In a field playback experiment, Gerhardt et al. [40] showed that male subjects "add up" quacks from two calls coming from different locations separated by 180° and respond with calls having more than the correct number of quacks (Figure 3a). Studies of sequential integration in Cope's gray treefrogs, *H. chrysoscelis* [41], and the eastern gray treefrog, *H. versicolor* [42], have found that females will group interleaved sequences of pulses over similarly large angles of separation. Farris et al. [43, 44] and Farris and Ryan [45] have shown that female túngara frogs, *Engystomops* (formerly *Physalaemus*) *pustulosus*, are also permissive of spatial separation in grouping the sequential "whine" and "chuck" elements of the two-part complex advertisement call (Figure 3b). Females grouped these two elements even when they were separated by up to 135°. Grouping based on spatial origin became considerably *less* permissive, however, when females were able to choose to group the whine with either of two spatially separated chucks, more often choosing the chuck located closest to the whine [45].

Recent work by Farris and Ryan [45] on túngara frogs has investigated the temporal sequence of complex calls as a possible "schema-based" [4] cue for sequential integration. In this species, males are morphologically constrained to produce a "whine" followed by one or more "chucks" [45], so the normal call sequence is fixed. In a clever playback experiment, females were given a choice between grouping whines with chucks in either correct or incorrect temporal positions relative to the whine. The results indicated that females were more likely to group whines with chucks that occurred in the correct temporal positions. Processing this temporal sequence cue may involve top-down mechanisms based on the listener's evolved expectation of hearing the natural sequence of a whine followed by chucks [45].

As the studies highlighted in this and the previous section illustrate, we are just beginning to understand the auditory grouping rules that anurans use in the context of acoustic signaling. The salience of some well-studied cues (e.g., common F_0 and common spatial origin)

remains uncertain and could well vary among species. In addition, the salience of various cues may differ among tasks requiring sequential versus simultaneous integration (as is sometimes the case in humans). Our ability to evaluate the salience of various grouping cues in frogs may also depend on the experimental task (e.g., no-choice versus choice) or the methodological approach (e.g., reflex modification versus playback experiment). These issues are worth bearing in mind in designing future studies of auditory grouping in frogs.

Auditory streaming

Numerous studies of sequential integration in humans [6, 8] and other animals [9] have investigated a phenomenon referred to as "auditory streaming." This term describes our ability to perceptually segregate interleaved or overlapping sequences of sounds into separate streams that can be selectively attended and followed through time. Studies of auditory streaming commonly employ stimuli composed of repeating sequences of two interleaved sounds (A and B) differing in frequency or some other acoustic property (e.g., ABABABAB…). For example, when the two rapidly repeating tones are similar in frequency, we tend to hear a coherent tone sequence jumping up and down in frequency. However, when the frequency difference (ΔF) between the two tones is large enough, our percept changes and we tend to hear two separate sequences, or "streams," corresponding to tones of different frequencies (e.g., A–A–A–A–… and –B–B–B–B…). Work on auditory streaming is important because it investigates the auditory system's ability to integrate temporally separated sounds elements into coherent auditory streams that unfold through time even when other intervening sounds are present.

One recent study investigated whether frogs might exploit ΔF as a cue for segregating streams of conspecific from heterospecific calls. Inspired by studies of auditory streaming in humans based on interleaved tone sequences (e.g., ABABAB…), Nityananda and Bee [46] used interleaved pulse sequences to investigate streaming in Cope's gray treefrogs, *H. chrysoscelis.* They manipulated the attractiveness of a discrete series of "target" pulses (simulating a conspecific call) by presenting it in such a way that the target pulses were interleaved with a continuous sequence of unattractive "distractor" pulses (simulating an overlapping heterospecific call, such as that of an American toad, *Anaxyrus americanus*). On different trials, the target and distractor pulses had the same or different frequencies. Recognition of the target signal increased as an increasing function of ΔF and in a way reminiscent of the improvements in word and melody recognition seen in similar studies of humans [4, 47]. Significant improvements in performance were observed at ΔFs simulating differences among species, but not at smaller differences more similar to among-individual differences within gray treefrogs.

Studies of the neural basis of auditory streaming in birds and mammals [8, 9] have implicated frequency selectivity, forward masking, and long-term adaptation as contributing to stream segregation based on frequency differences. All of these processes are also known from the anuran auditory system [48–50], though they have not yet been investigated in the context of auditory streaming in frogs. The contribution of these and other neural processes (e.g., neural pulse "counting" [31]) to auditory streaming in frogs is a question that deserves additional research. Other open questions concern the extent to which frogs may exploit cues other than frequency differences or combinations of cues, and their ability to segregate within-species (i.e., among-individual) differences in various cues.

Auditory induction

In acoustically and structurally complex environments, sounds of interest may often be masked by brief, intermittent loud sounds (e.g., a cough, a clap of thunder, a door shutting). Humans [51], nonhuman primates [52], and songbirds [53] possess the remarkable ability to

reconstruct percepts of sounds masked by (or even *replaced* by) brief bursts of noise. Known as "auditory induction" [51], this ability is responsible for "phonemic restoration" in humans and underlies the so-called "continuity illusion." Recent studies of frogs, however, suggest that auditory induction may not be a general feature of sound source perception in vertebrates. Studies of Cope's gray treefrogs, *H. chrysoscelis* [54], eastern gray treefrogs, *H. versicolor* [55], and túngara frogs, *E. pustulosus* [Baugh AT, Bee MA, and Ryan MJ unpublished data] have so far failed to uncover evidence that frogs perceptually reconstruct segments of advertisement calls replaced by noise bursts even though humans listening to the same test stimuli are able to do so (Figure 4). In contrast to birds and mammals, current evidence, therefore, suggests that the frog auditory system may be incapable of auditory induction. The mechanistic basis for these apparent taxonomic differences is an issue that requires further investigation. Additional studies of auditory induction in other frogs are also needed.

Conclusions

Research on sound source perception in anurans represents an exciting frontier in the neuroethological study of acoustic communication in these animals. As the behavioral studies reviewed here make clear, there are some fascinating parallels and notable differences in sound source perception between frogs and other animals, as well as among different species of frogs. Demonstrated similarities and differences within and among taxa create the solid foundation necessary for comparative neuroethological studies that can better inform us about the evolution and function of underlying neural mechanisms. Future work aimed at uncovering commonalities and diversity in the neural mechanisms of sound source perception in frogs and other animals are sure to be worthwhile and informative.

Acknowledgments

References

1. Hulse SH. Auditory scene analysis in animal communication. Advances in the Study of Behavior. 2002; 31:163–200.

2. Bee MA, Micheyl C. The cocktail party problem: What is it? How can it be solved? And why should animal behaviorists study it? Journal of Comparative Psychology. 2008; 122:235–51. [PubMed: 18729652]

3•. McDermott JH. The cocktail party problem. Current Biology. 2009; 19:R1024–R7. This short paper is a noteworthy review of research on the "cocktail party problem" in humans. [PubMed: 19948136]

4. Bregman, AS. Auditory Scene Analysis: The Perceptual Organization of Sound. MIT Press; 1990.

5. Darwin, CJ.; Carlyon, RP. Auditory grouping. In: Moore, BCJ., editor. Hearing. Academic Press; 1995. p. 387-424.

6. Carlyon RP. How the brain separates sounds. Trends Cogn Sci. 2004; 8:465–71. [PubMed: 15450511]

7. Yost, WA.; Popper, AN.; Fay, RR. Auditory Perception of Sound Sources. Springer; 2008.

8•. Shamma SA, Micheyl C. Behind the scenes of auditory perception. Curr Opin Neurobiol. 2010; 20:361–6. This paper reviews recent findings on the neurophysiological basis of auditory scene analysis, with an emphasis on auditory streaming. [PubMed: 20456940]

9. Fay, RR. Sound source perception and stream segregation in nonhuman vetebrate animals. In: Yost, WA.; Popper, AN.; Fay, RR., editors. Auditory Perception of Sound Sources. Springer; 2008. p. 307-23.

10. Webster, DB.; Fay, RR.; Popper, AN. The Evolutionary Biology of Hearing. New York: Springer; 1992.

11. Christensen-Dalsgaard J, Carr CE. Evolution of a sensory novelty: tympanic ears and the associated neural processing. Brain Research Bulletin. 2008; 75:365–70. [PubMed: 18331899]

12••. Schnupp JWH, Carr CE. On hearing with more than one ear: Lessons from evolution. Nature Neuroscience. 2009; 12:692–7. This article reviews similarities and differences in how birds and mammals process binaural cues for sound localization. The subject matter should be of general interest to neuroethologists interested in hearing and sound communication as it nicely illustrates the potential for evolution to find different solutions to common problems in different phylogenetic lineages.

13. Gerhardt, HC.; Huber, F. Acoustic Communication in Insects and Anurans: Common Problems and Diverse Solutions. Chicago University Press; 2002.

14. Narins, PM.; Feng, AS.; Fay, RR.; Popper, AN. Hearing and Sound Communication in Amphibians. Springer; 2007.

15. Gerhardt HC. Sound pressure levels and radiation patterns of vocalizations of some North American frogs and toads. Journal of Comparative Physiology. 1975; 102:1–12.

16. Narins PM. Effects of masking noise on evoked calling in the Puerto Rican coqui (Anura, Leptodactylidae). Journal of Comparative Physiology. 1982; 147:439–46.

17. Arak A. Sexual selection by male-male competition in natterjack toad choruses. Nature. 1983; 306:261–2.

18••. Vélez, A.; Schwartz, JJ.; Bee, MA. Anuran signal perception in noisy environments. In: Brumm, H., editor. Animal Communication and Noise. Springer; in press. This book chapter reviews behavioral and physiological studies of sound source perception in more detail than is possible in the present short review. The two-part focus of the chapter describes both the magntiude of the frog's "cocktail party problem" from a functional perspective and the mechanisms by which frog listeners attempt to overcome this problem. This chapter also more thoroughly relates research on frog communication and sound source perception to relevant studies of humans and other animals

19••. Schwartz, JJ.; Bee, MA. Anuran signal production in noisy environments. In: Brumm, H., editor. Animal Communication and Noise. Springer; in press. In this book chapter, the authors review the extensive literature on how the vocal signaling behaviors of male frogs are adapted to ameliorate auditory masking and call interference in the acoustic environment of a breeding chorus. An important take-home message from this review is that the demands of sound source perception in acoustically complex environments place signalers under strong selection to signal in ways that reduce the magnitude of cocktail-party-like problems in animal communication systems

20. Bronkhorst AW. The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions. Acustica. 2000; 86:117–28.

21. Schwartz JJ, Gerhardt HC. Spatially mediated release from auditory masking in an anuran amphibian. Journal of Comparative Physiology A. 1989; 166:37–41.

22. Bee MA. Sound source segregation in grey treefrogs: Spatial release from masking by the sound of a chorus. Animal Behaviour. 2007; 74:549–58.

23•. Nityananda V, Bee MA. Spatial release from masking in a free-field source identification task by gray treefrogs. Hearing Research. in press. In this study, target stimuli were designed to investigate spatial release from masking as a function of the spectral content of the signal. On various trials, the target had either the normal bimodal spectrum (with peaks at 1.3 kHz and 2.6 kHz) or a modified unimodal spectrum comprising only one of the two spectral peaks of the bimodal spectrum. In this species, the lower and upper spectral peaks are primarily encoded by separate inner ear organs, the amphibian papilla and the basilar papilla, respectively. Somewhat surpisingly, the magnitude of spatial release was largely independent of the spectral content of the signal.

24. Bee MA. Finding a mate at a cocktail party: Spatial release from masking improves acoustic mate recognition in grey treefrogs. Animal Behaviour. 2008; 75:1781–91. [PubMed: 19412318]

25. Ratnam R, Feng AS. Detection of auditory signals by frog inferior collicular neurons in the presence of spatially separated noise. J Neurophysiol. 1998; 80:2848–59. [PubMed: 9862889]

26. Lin WY, Feng AS. Free-field unmasking response characteristics of frog auditory nerve fibers: comparison with the responses of midbrain auditory neurons. Journal of Comparative Physiology A. 2001; 187:699–712.

27. Lin WY, Feng AS. GABA is involved in spatial unmasking in the frog auditory midbrain. Journal of Neuroscience. 2003; 23:8143–51. [PubMed: 12954877]

28. Füllgrabe C, Berthommier F, Lorenzi C. Masking release for consonant features in temporally fluctuating background noise. Hearing Research. 2006; 211:74–84. [PubMed: 16289579]

29•. Vélez A, Bee MA. Dip listening and the cocktail party problem in grey treefrogs: signal recognition in temporally fluctuating noise. Animal Behaviour. 2011; 82:1319–27. This study measured "signal recognition thresholds," which are conceptually analogous to the "speech reception thresholds" measured in some studies of sound source perception in humans [20, 28]. The null hypothesis was that temporal fluctuations in the level of a masker have no effect on signal recognition thresholds. This null was tested against the alternative hypotheses that masker fluctuations result in either lower recognition thresholds (i.e., "dip listening") or higher thresholds (i.e., "modulation masking"). Dip listening was supported at low modulation rates (< 5 Hz), whereas modulation masking was found at a fast rate more similar to the pulse rate of the target signal (e.g., 40 Hz). Subsequent analyses of natural chorus sounds indicated that low-frequency "dips" in amplitude were free from the faster rates of modulation characteristic of the pulse rate, suggesting that under natural listening conditions, dip listening might be important in a receiver's perception of male advertisement calls. [PubMed: 22389519]

30•. Vélez A, Bee MA. Signal recognition by frogs in the presence of temporally fluctuating chorus-shaped noise. Behavioral Ecology and Sociobiology. 2010; 64:1695–709. This study is significant because it includes an analysis of the temporal structure of the sounds of choruses of several different species of frogs. These analyses revealed some commonalities and some species differences. In all cases, the modulation spectra of natural chorus sounds exhibited prominent low-frequency (< 5–10 Hz) temporal fluctuations. These fluctuations were attributed to both the timing of discrete calls produced by individual males in the chorus and the possible imposition of low-frequency modulations by enviornmental factors (e.g., air turbulence). In species with pulsatile calls, the pulse rate was also observed as a distinct peak in the modulation spectrum. [PubMed: 21170157]

31. Edwards CJ, Alder TB, Rose GJ. Auditory midbrain neurons that count. Nature Neuroscience. 2002; 5:934–6.

32•. Bee MA, Vélez A, Forester JD. Sound level discrimination by gray treefrogs in the presence and absence of chorus-shaped noise. Journal of the Acoustical Society of America. in review. In this study, females of Cope's gray treefrogs were given a choice between two identical stimuli that differed only by 2 dB, 4 dB, or 6 dB in overall level. Tests were conducted in a fully factorial design at two nominal signal levels (79 dB or 85 dB for the unattenuated signal) and in the presence and absence of noise simulating breedng choruses. At the 79-dB nominal level, females exhibit robust preferences for signals that were more intense (by 2 dB, 4 dB, and 6 dB) than the alternative in both the presence and absence of noise. At the 85-dB nominal signal level, discrimination in favor of the more intense signal was observed at 4 dB and 6 dB in both the presence and absence of noise.

33. Darwin, CJ. Spatial hearirng and perceiving sources. In: Yost, WA.; Popper, AN.; Fay, RR., editors. Auditory Perception of Sound Sources. Springer; 2008. p. 215-32.

34. Simmons AM, Bean ME. Perception of mistuned harmonics in complex sounds by the bullfrog (*Rana catesbeiana*). Journal of Comparative Psychology. 2000; 114:167–73. [PubMed: 10890588]

35. Simmons AM. Selectivity for harmonic structure in complex sounds by the green treefrog (*Hyla cinerea*). Journal of Comparative Physiology A. 1988; 162:397–403.

36. Gerhardt HC, Allan S, Schwartz JJ. Female green treefrogs (*Hyla cinerea*) do not selectively respond to signals with a harmonic structure in noise. Journal of Comparative Physiology A. 1990; 166:791–4.

37. Simmons AM, Buxbaum RC, Mirin MP. Perception of complex sounds by the green treefrog, *Hyla cinerea*: Envelope and fine-structure cues. Journal of Comparative Physiology A. 1993; 173:321–7.

38. Bodnar DA. The separate and combined effects of harmonic structure, phase, and FM on female preferences in the barking treefrog (*Hyla gratiosa*). Journal of Comparative Physiology A. 1996; 178:173–82.

39•. Bee MA. Spectral preferences and the role of spatial coherence in simultaneous integration in gray treefrogs (*Hyla chrysoscelis*). Journal of Comparative Psychology. 2010:124. As the first study of spatially mediated simultaneous integration in frogs, the results of this study were somehwat surprising. In contrast to earlier studies of sequential itnegration based on common spatial origin [40–45], the results of this study indicated that even a small amount of spatial incoherence was enough to interfere with call perception.

40. Gerhardt HC, Roberts JD, Bee MA, Schwartz JJ. Call matching in the quacking frog (*Crinia georgiana*). Behavioral Ecology and Sociobiology. 2000; 48:243–51.

41. Bee MA, Riemersma KK. Does common spatial origin promote the auditory grouping of temporally separated signal elements in grey treefrogs? Animal Behaviour. 2008; 76:831–43. [PubMed: 19727419]

42. 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.

43. Farris HE, Rand AS, Ryan MJ. The effects of spatially separated call components on phonotaxis in túngara frogs: evidence for auditory grouping. Brain Behavior and Evolution. 2002; 60:181–8.

44. Farris HE, Rand AS, Ryan MJ. The effects of time, space and spectrum on auditory grouping in túngara frogs. Journal of Comparative Physiology A. 2005; 191:1173–83.

45••. Farris HE, Ryan MJ. Relative comparisons of call parameters enable auditory grouping in frogs. Nat Commun. 2011; 2:410. This study is notworthy for two important reasons. First, it investigated *relative* comparisons in the context of auditory grouping. In so doing, this study is the first to show that grouping based on common spatial origin can be more precise when receivers are forced to choose precisely which sound elements get grouped together over time. Second, this study is the first to show that the natural order of a temporally structured call can itself be exploited as a "schema based" auditory grouping cue. [PubMed: 21811239]

46••. 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. The experiments described in this study were designed to incorporate elements of previous experiments on the role of frequency differences in auditory streaming of tone sequences and concurrent voices in humans. The target and distractor pulses were designed to simulate the voices of a conspecific and heterospecific frog calling at the same time. The pulses of the target and distractor were interleaved both to simulate the common use of interleaved tone sequences in studies of auditory streaming (e.g., ABABAB…) and to exploit the robust preferences females exhibit for signals with the correct pulse rate. The reasoning behind this manipulation was that, if females perceived the target and distractor as a single stream, then it would have an unattractive pulse rate and fail to elicit phonotaxis. If, on the other hand, females were able to segregate the target from the distractor, then they should perceive the attractive pulse rate of the target as distinct from the pulsatile distractor, and hence approach the signal. Results showed that recognition of the target was an increasing function of frequency separation from the distractor. [PubMed: 21698268]

47. Bird, J.; Darwin, CJ. Effects of a difference in fundamental frequency in separating two sentences. In: Palmer, AR.; Rees, A.; Summerfield, AQ.; Meddis, R., editors. Psychophysical and Physiological Advances in Hearing. Whurr; 1998. p. 263-9.

48. Hillery CM, Fay RR. Forward masking and suppression in the midbrain of the southern grey treefrog (*Hyla chrysoscelis*). Journal of Comparative Physiology. 1982; 146:435–47.

49. Megela AL, Capranica RR. A Neural and behavioral study of auditory habituation in the bullfrog, *Rana catesbeiana*. Journal of Comparative Physiology. 1983; 151:423–34.

50. Zakon, HH.; Wilczynski, W. The physiology of the anuran eighth nerve. In: Fritzsch, B.; Wolkowiak, W.; Ryan, MJ.; Wilczynski, W.; Hetherington, T., editors. The Evolution of the Amphibian Auditory System. Wiley; 1988. p. 125-55.

51. King AJ. Auditory neuroscience: Filling in the gaps. Current Biology. 2007; 17:R799–R801. [PubMed: 17878047]

52. Petkov CI, O'Connor KN, Sutter ML. Illusory sound perception in macaque monkeys. Journal of Neuroscience. 2003; 23:9155–61. [PubMed: 14534249]

53. Seeba F, Klump GM. Stimulus familiarity affects perceptual restoration in the European starling (*Sturnus vulgaris*). PLoS ONE. 2009; 4:e5974. [PubMed: 19551146]

54•. Seeba F, Schwartz JJ, Bee MA. Testing an auditory illusion in frogs: perceptual restoration or sensory bias? Animal Behaviour. 2010; 79:1317–28. In this study (and also that by Schwartz et al. [55]), the perception of illusory pulses was investigated by exploiting the robust directional preferences of females for longer calls with more pulses. In the key manipulation, females were given a choice between two alternatives in which the number of actual pulses was similar, but the longer alternative had gaps inserted and filled with noise, making its total duration longer. If females perceived illusory pulses, they were predicted to choose the longer alternative. But this was not the case. Females responded to the two alternatives as if they had equal number of pulses, thus refuting the hypothesis that females experienced illusory pulses. [PubMed: 20514342]

55. 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.

56. Bee MA, Schwartz JJ. Behavioral measures of signal recognition thresholds in frogs in the presence and absence of chorus-shaped noise. Journal of the Acoustical Society of America. 2009; 126:2788–801. [PubMed: 19894854]

Highlights

- Sound source perception in animal communication is not well understood.

- Frogs solve a biological analogue of the "cocktail party problem" to communicate.

- This article reviews studies of how they do so using a limited number of acoustic cues.

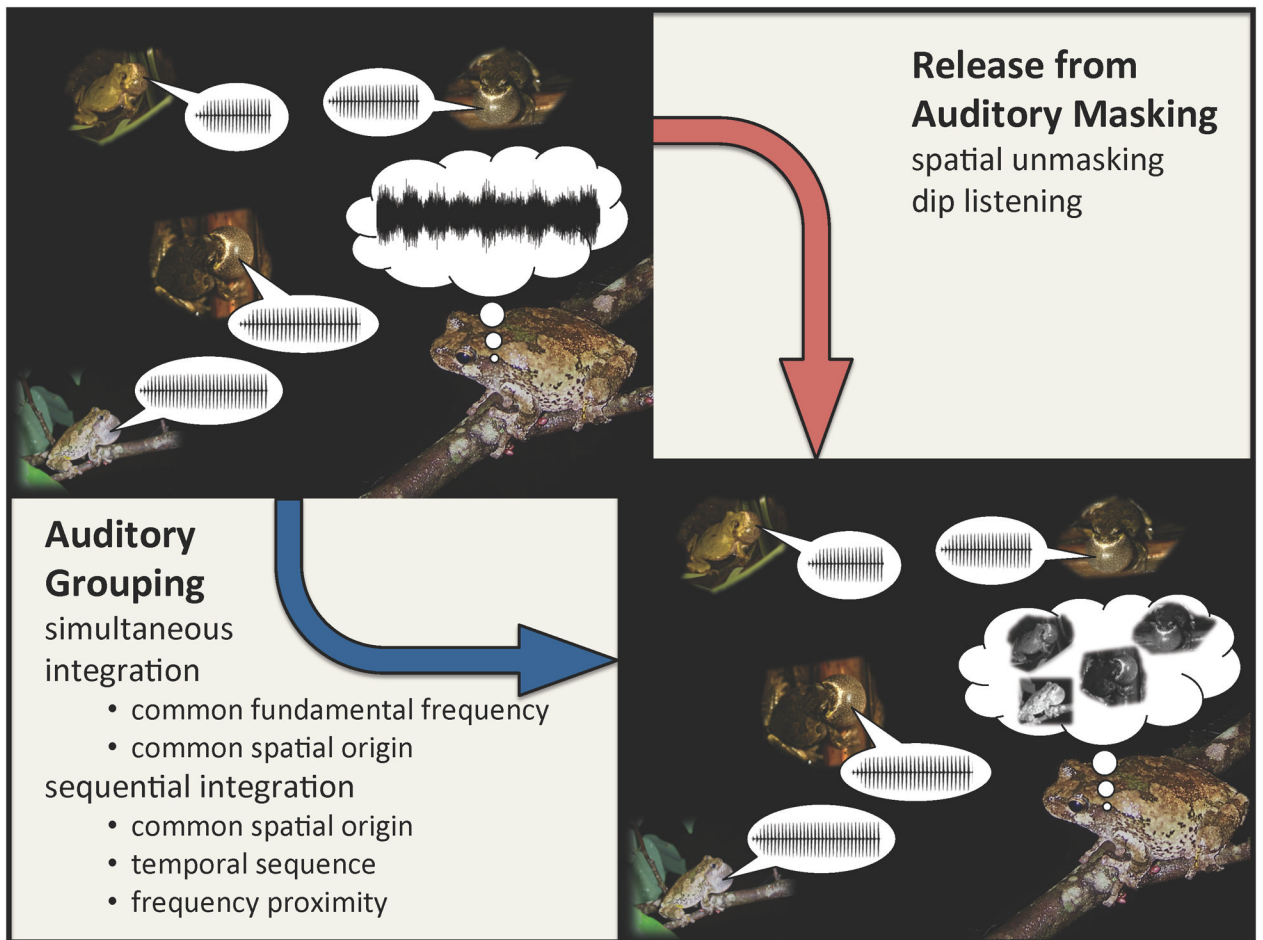- The focus is on mechanisms of masking release and auditory grouping.

Figure 1.
Two components of sound source perception in frogs. The basic auditory problem that frogs encounter in natural social aggregations (depicted in the upper left) is that all of the sound pressure waves produced by individual calling males in a chorus add together to form a single, composite sound pressure waveform. This is a general problem in hearing, and it is a specific biological analogue of the human "cocktail party problem" [2]. It is the composite waveform that impinges on the frog's body and sets the tympana into motion. In order to make sense of the complex acoustic scenes typical of breeding choruses, a listening frog's auditory system must decompose this composite waveform to create coherent percepts that correspond to individual calling males as distinct sound sources in the environment (depicted in the lower right). Two challenges emerge from this general problem, one related to auditory masking and the other related to auditory grouping. Behavioral and physiological studies of frogs indicate they may overcome these two challenges by exploiting some of the same features of the acoustic scene that enable humans to understand speech under cocktail-party-like listening conditions. Frogs can experience significant release from auditory masking by exploiting spatial separation between signals and noise and by "listening in the dips" of the natural fluctuations that are present in the level of natural chorus noise. Frogs also appear to integrate simultaneous spectral components and temporally separated sound elements using well-known auditory grouping cues related to commonalities shared by the sound elements produced by a single source.
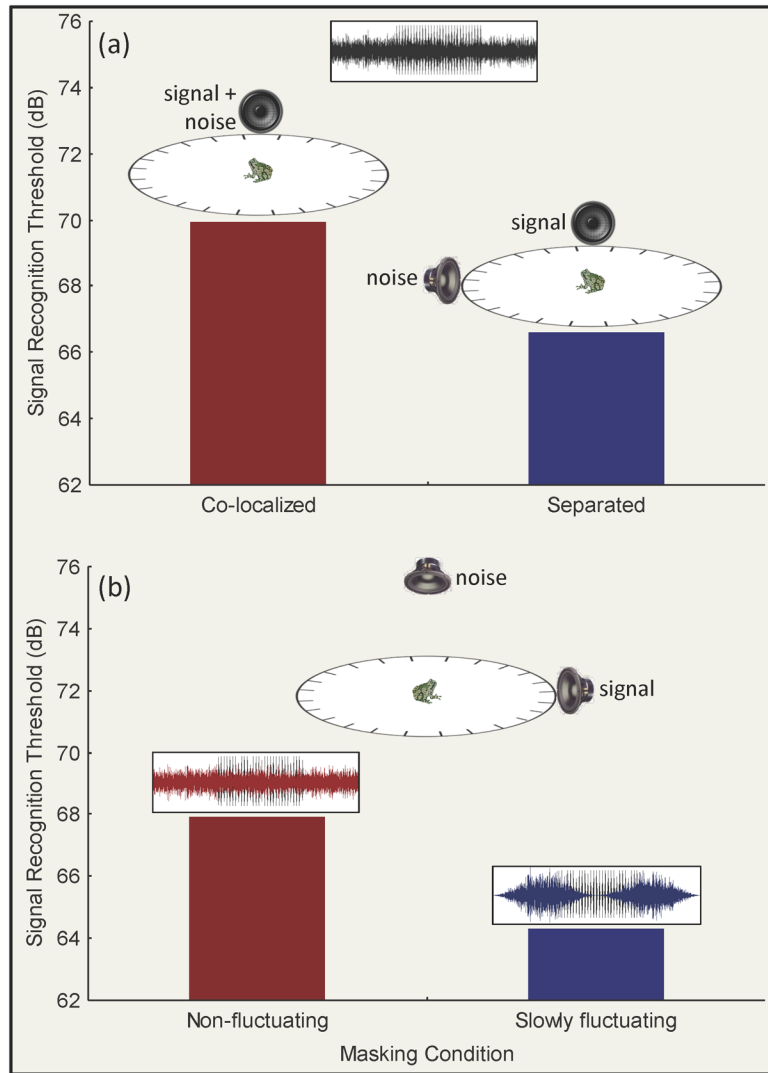
Figure 2.
Release from auditory masking in frogs. In the social environment of a breeding chorus, frogs must often contend with high levels of background noise that potentially masks signals of interest. Recent behavioral studies of female Cope's gray treefrogs (*Hyla chrysoscelis*) have used adaptive tracking procedures to investigate release from auditory masking in the presence of chorus-like noise by measuring "signal recognition thresholds" (defined as the lowest signal level that elicits phonotaxis) [23, 29, 56] under various listening conditions. **(a)** When conspecific calls (signals) and chorus-like noise were separated by 90° around a 2-m-diameter circular test arena, signal recognition thresholds were about 4 dB lower compared to conditions in which the same signals and noise were co-localized [23]. **(b)** Subjects also experienced about 4 dB of masking release when they listened for pulsed advertisement calls presented in chorus-like noise that slowly fluctuating (e.g., sinusoidally amplitude modulated (SAM) noise; 1.25 Hz) compared to when listening in a non-fluctuating, "flat" noise condition [29]. Together, results from such studies indicate that the frog auditory system can exploit spatial separation between signals and noise and temporal fluctuations in background noise levels in recognizing conspecific calls. (Note: Frogs, speakers, and test arenas not drawn to scale.)
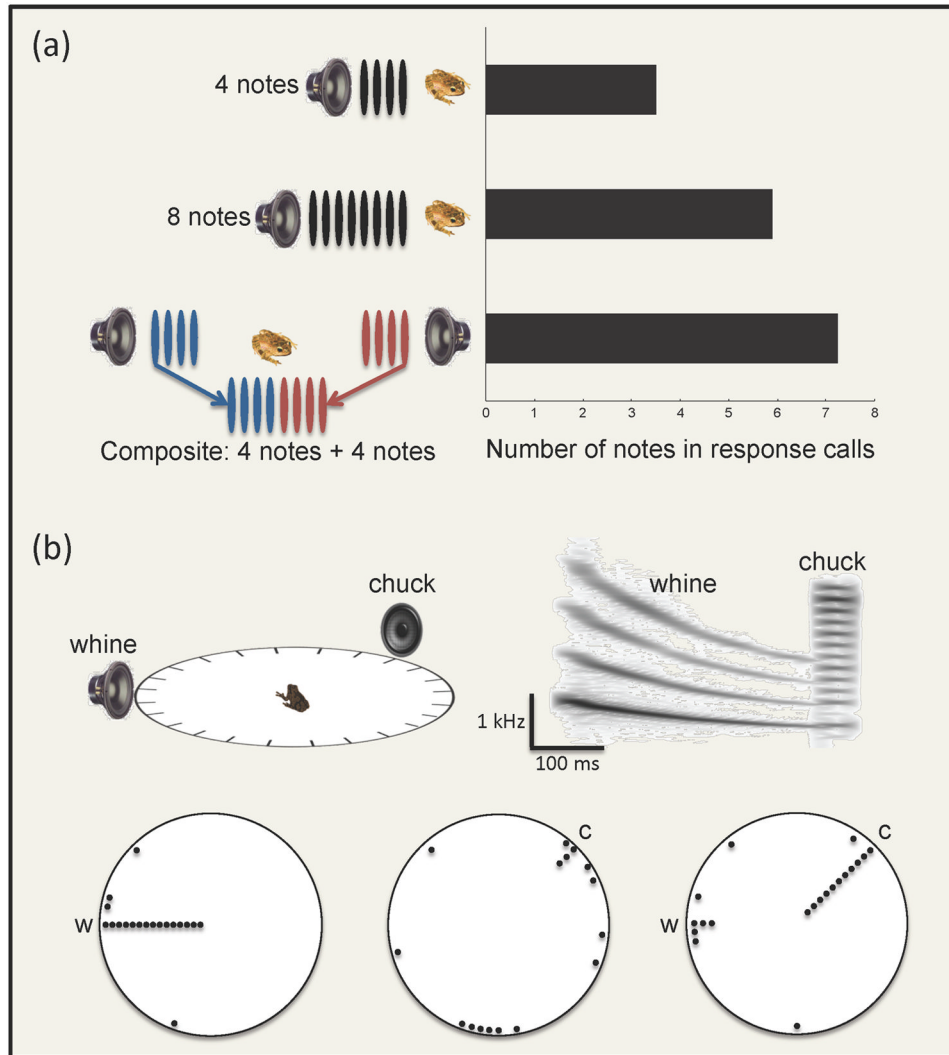
Figure 3.
Sequential integration in frogs based on common spatial origin. Frog calls typically comprise multiple sound elements (e.g., notes or pulses) separated in time. For many species, recognizing conspecific calls or selecting high-quality mates requires that receivers perceptually bind temporally separated sound elements into coherent "auditory objects." Several studies directly or indirectly investigating the role of common spatial origin as an auditory grouping cue indicate that frogs may be fairly permissive of spatial incoherence in recognizing or preferring certain calls. **(a)** In a field playback experiment, males of the Australian quacking frog (*Crinia georgiana*) were shown to closely match the number of notes in stimulus calls simulating nearby neighbors [40]. However, males gave calls with similar numbers of response notes to 8-note stimulus calls coming from one direction and composite 8-note calls in which the first and second four notes of the call came from speakers separated by 180° [40]. **(b)** Laboratory studies of phonotaxis in túngara frogs (*Engystomops pustulosus*) have also found that females can be permissive of wide angular separations between the preferred "whine" and "chuck" components of complex calls [45]. Whines (w) presented alone elicited robust phonotaxis (indicated here by orientation angles at which females exited the circular test arena relative to the position of the speaker broadcasting the stimulus [45]. Chucks (c) alone failed to elicit phonotaxis; however, whine-

chuck sequences elicited phonotaxis directed toward the chuck even when it was separated from the whine by up to 135° [45]. Together, these and other studies [41–44] indicate that frogs can be permissive of spatial incoherence and will perceptually group sounds together even though they originate from different locations.
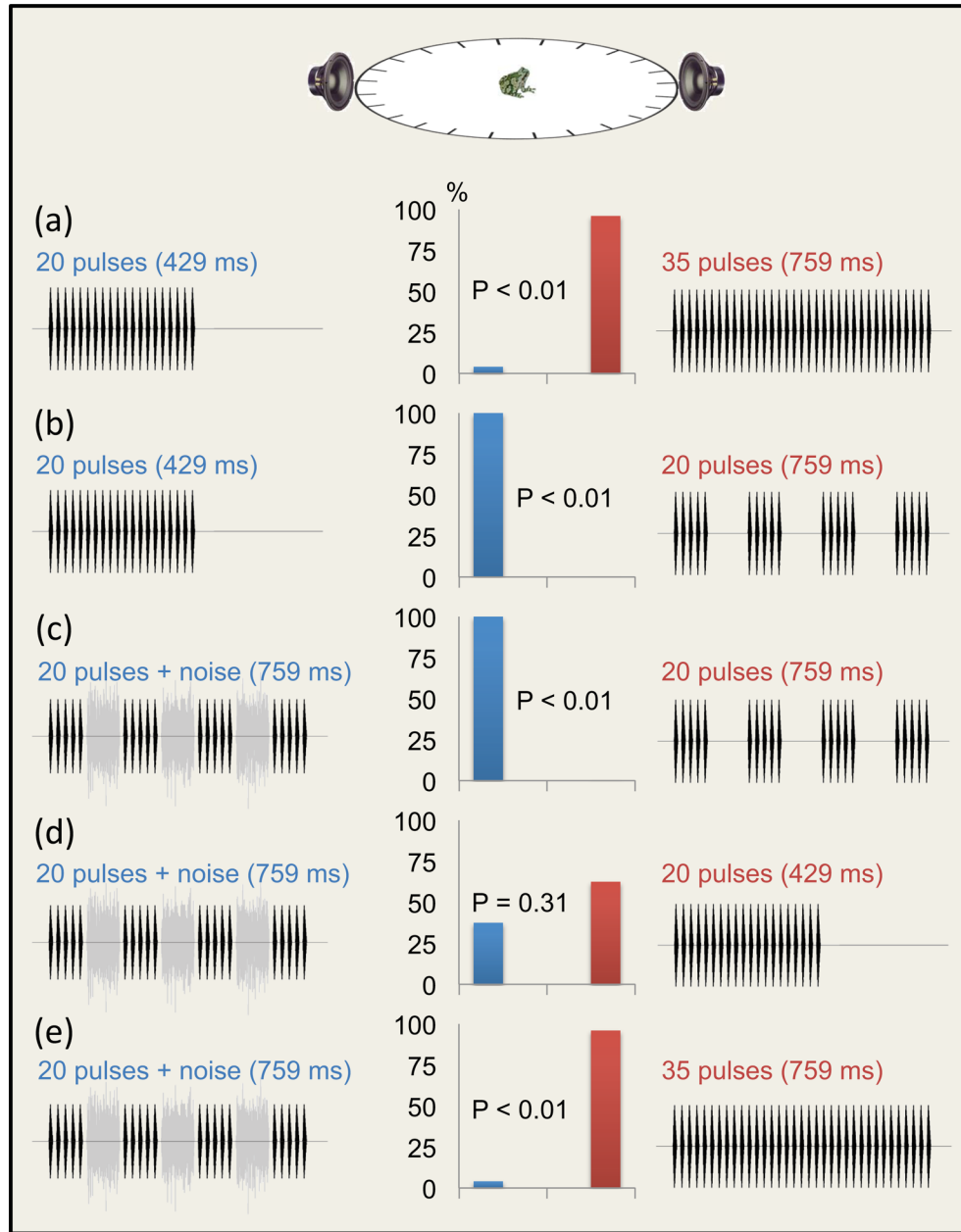
Figure 4.
Can frogs fill in the gaps? Studies of auditory induction in frogs have, thus far, failed to uncover convincing evidence that frogs experience a biological analogue of "phonemic restoration" in hearing illusory signal elements (e.g., pulses) that have been replaced by broadband noise. Auditory induction has been investigated in two species of treefrog using two-choice discrimination tests that exploited female preferences for longer signals that contain more pulses [54, 55]. **(a)** Females of Cope's gray treefrog (*H. chrysoscelis*) [55] prefer relatively longer calls composed of continuous trains of pulses. **(b)** Introducing short gaps into preferred calls by removing pulses can shift female preferences to shorter but nevertheless continuous pulse trains. **(c)** Filling these gaps with noise can restore the relative attractiveness of the longer signal when compared to equivalent-duration signals with gaps. **(d, e)** However, there is no evidence that restored attractiveness results because females hear

continuous but illusory trains of pulses; if this were the case, we would expect both **(d)** females to prefer longer, "gap-filled" calls over shorter trains of continuous pulses (which they do not) and **(e)** gap-filled calls and continuous pulse trains of equivalent duration to be similarly attractive (which they are not).