Behind the mask(ing): How frogs cope with noise

Norman Lee[1] ● Alejandro Vélez[2] ● Mark Bee[3,4]

[1] Department of Biology, St. Olaf College, 1520 St. Olaf Ave, Northfield, MN, 55057, USA
e-mail: lee33@stolaf.edu
ORCID: 0000-0001-6198-710X

[2] Department of Biology, San Francisco State University, 1600 Holloway Ave, San Francisco, CA 94132, USA
e-mail: avelezm@sfsu.edu
ORCID: 0000-0002-7636-8405

[3] Department of Ecology, Evolution, and Behavior, University of Minnesota – Twin Cities, 1479 Gortner Ave, St. Paul, MN 55108, USA
[4] Graduate Program in Neuroscience, University of Minnesota – Twin Cities, 321 Church Street SE, Minneapolis, MN 55455, USA
e-mail: mbee@umn.edu
ORCID: 0000-0002-6770-9730

* Corresponding author: Norman Lee, lee33@stolaf.edu

Abstract

Albert Feng was a pioneer in the field of auditory neuroethology who used frogs to investigate the neural basis of spectral and temporal processing and directional hearing. Among his many contributions was connecting neural mechanisms for sound pattern recognition and localization to the problems of auditory masking that frogs encounter when communicating in noisy, real-world environments. Feng's neurophysiological studies of auditory processing foreshadowed and inspired subsequent behavioral investigations of auditory masking in frogs. For frogs, vocal communication frequently occurs in breeding choruses, where males form dense aggregations and produce loud species specific advertisement calls to attract potential mates and repel competitive rivals. In this review, we aim to highlight how Feng's research advanced our understanding of how frogs cope with noise. We structure our narrative around three themes woven throughout Feng's research – spectral, temporal, and directional processing – to illustrate how frogs can mitigate problems of auditory masking by exploiting frequency separation between signals and noise, temporal fluctuations in noise amplitude, and spatial separation between signals and noise. We conclude by proposing future research that would build on Feng's considerable legacy to advance our understanding of hearing and sound communication in frogs and other vertebrates.

Keywords Auditory scene analysis Comodulation masking release Energetic masking Matched filtering Spatial release from masking

Introduction

Natural acoustic environments are often complex, consisting of multiple sounds that overlap in frequency and time and that are produced by multiple different sources (Bregman 1990; Yost 2008). These sound sources can be biotic, which may include conspecific or heterospecific communication signals, or abiotic such as environmental or anthropogenic noise (Brumm and Slabbekoorn 2005; Brumm 2013; Wiley 2015). Listening to a specific source in a complex, multi-source environment presents significant challenges because overlapping sounds and background noise can make it more difficult or even impossible to hear signals of interest, a phenomenon known generally as "auditory masking" (Patterson and Green 2012). For humans, large social gatherings, such as a crowded restaurant or a cocktail party, are one example of a complex acoustic environment where following one conversation among many is made more difficult due to auditory masking. This difficulty is aptly termed the "cocktail party problem" (Cherry 1953; Bronkhorst 2000; McDermott 2009; Middlebrooks et al. 2017). As a general rule, auditory masking is most pronounced when competing sounds are similar in frequency, occur at the same time, and originate from the same location. Intensive studies of human hearing and speech perception have revealed a number of mechanisms that enable us to cope with problems of auditory masking (Moore 2013). Many of these mechanisms, however, are not unique to human hearing (Fay and Popper 2000).

In this article, we honor the legacy of Albert Feng by reviewing research he inspired on how frogs cope with problems of auditory masking (Feng and Ratnam 2000; Feng and Schul 2007). Like humans, frogs also communicate in noisy social gatherings where they experience problems analogous to the human cocktail party problem. During their breeding season, many frog species form large social aggregations where males call loudly to attract receptive females and repel rival males (Gerhardt and Huber 2002). These aggregations may consist of hundreds of males, often of multiple species, densely packed into aquatic habitats that are suitable for breeding. An individual male's advertisement calls can reach sound pressure levels (SPLs) as high as 90 to 110 dB SPL (re 20 μPa) at a distance of 1 m (Gerhardt 1975), and the sustained background noise level in choruses can easily reach or exceed 80 dB to 90 dB SPL (Narins 1982; Halfwerk et al. 2016; Tanner and Bee 2019). Choruses can be audible to humans from distances of up to 2 km away from the breeding site (Arak 1983). In the environment of a breeding chorus, high levels of background noise and overlapping signals lead to auditory masking that is manifest as impaired signal detection, recognition, discrimination, and localization (Bee 2012, 2015; Vélez et al. 2013b).

How do frogs cope with the noise they encounter in breeding choruses? Albert Feng took up this question in two important reviews on sound processing in "real-world" environments (Feng and Ratnam 2000; Feng and Schul 2007). He and his co-authors described previous studies on the neural processing of spectral, temporal, and directional information in frogs and discussed (and in some cases speculated) how such processes might contribute to hearing the calls of individual males amid the cacophony of a chorus. It was also recognized, however, that the sensory basis of hearing in complex, multi-source environments was inadequately understood, and that more studies that employed psychoacoustic techniques were needed to further elucidate the perceptual abilities of auditory systems, and the processes and cues that auditory systems utilize in solving the cocktail-party-like problems (Feng and Ratnam 2000; Feng and Schul 2007). In the years following Feng's two reviews, we have learned a great deal more about the magnitude of the frog's cocktail-party-like problem and how they cope with it. Here, we review much of this more recent work by reflecting on how Feng's neurophysiological studies of spectral processing, temporal processing, and directional hearing foreshadowed and inspired subsequent behavioral investigations of auditory masking in frogs.

Exploiting frequency differences between signals and noise

Matched spectral filtering

Amphibians are unique among vertebrates in having two distinct auditory organs in the inner ear tuned to different frequency ranges of airborne sound. Early work by Feng et al. (1975) demonstrated that auditory fibers innervating the amphibian papilla (AP) are tuned to low and intermediate sound frequencies (e.g., < 1.5 kHz), while auditory fibers innervating the basilar papilla (BP) are tuned to higher sound frequencies. The specific frequencies to which the AP and BP are tuned vary between species, but they are commonly found to be most sensitive to frequency peaks that are emphasized in conspecific advertisement calls (Gerhardt and Schwartz 2001). This match in the frequency sensitivity of the auditory periphery to the frequency content of conspecific calls inspired the “matched filter” hypothesis, which posits that the auditory periphery most optimally encodes the frequency content of conspecific advertisement calls, and filters out other sound frequencies (Capranica and Moffat 1983; Simmons 2013). One major theme of Feng’s research program was to discover how central auditory processes sharpen the matched spectral filtering that originates at the periphery. This body of work revealed how the simple, V-shaped tuning curves characteristic of auditory nerve fibers are transformed into much more complex, often bimodal tuning curves that become increasingly selective for frequencies emphasized in conspecific calls along the ascending auditory pathway, such that some neurons in the auditory midbrain and thalamus respond best, or only, to combinations of frequencies present in vocalizations and transduced by the AP and BP (Fuzessery and Feng 1981, 1982, 1983a, b; Gooler et al. 1993, 1996; Zhang and Feng 1998; Zhang et al. 1999; Goense and Feng 2005). A primary function of matched spectral filtering is that it allows a receiver to obtain “the highest signal-to-noise ratio in the frequency domain” for detecting conspecific vocalizations (p. 706; Capranica and Moffat 1983). Thus, matched filtering functions to improve the perception of conspecific vocalizations in the presence of noise, particularly noise that does not overlap the spectrum of conspecific vocalizations (e.g., the calls of other frog species in mixed-species choruses). The matched filter hypothesis predicts the co-evolution of spectral content in vocalizations and spectral tuning in the auditory system, a prediction generally well supported by comparative studies of frogs (Gerhardt and Schwartz 2001). While the vitality of the matched filter hypothesis has been questioned (Narins and Feng 2007), it seems clear that the frog’s peripheral and central auditory systems are adapted to exploit frequency differences between conspecific signals and noise at other frequencies.

A recent study of the American green treefrog (*Hyla cinerea*) extends the work of Feng and colleagues in describing a physical mechanism that functions to sharpen matched spectral filtering in an unexpected way, beginning at the tympanum itself (Lee et al. 2021). In this species, males produce advertisement calls with two spectral peaks centered at about 0.8-1.0 and 2.7-3.0 kHz (Gerhardt 1974; Lee et al. 2021) (Figs. 1a-b). These two spectral peaks are transduced primarily by the AP and BP, respectively, consistent with matched spectral filtering (Ehret and Capranica 1980). Audiograms based on both behavioral responses (Megela-Simmons et al. 1985) (Fig. 1d) and auditory evoked responses (Buerkle et al. 2014), as well as neural recordings from the auditory nerve and midbrain (Ehret and Capranica 1980; Miranda and

Wilczynski 2009a; Gall and Wilczynski 2015; Lee et al. 2017a) (Fig. 1d), demonstrate heightened sensitivity to the spectral peaks of the advertisement call (Figs. 1d-f). The auditory system is also most selective for frequencies at around 0.9 kHz and 3.0 kHz (Moss and Simmons 1986). Not surprisingly, females are generally most attracted to calls with both spectral peaks (Gerhardt 1981; Gerhardt and Höbel 2005; Lee et al. 2017a) (Fig. 1c), and combination sensitive neurons in the central auditory system also show selectivity for call-like sounds with both spectral peaks (Lee et al. 2017a) (Figs. 1e-f).

As in other frogs (Narins et al. 1988; Ehret et al. 1990; Jørgensen 1991; Jørgensen and Gerhardt 1991), sound can reach the internal surface of the green treefrog's tympana through the body wall and air-filled lungs via the glottis, mouth cavity, and Eustachian tubes. Lee et al. (2021) discovered that this lung-to-ear sound transmission pathway, unique to amphibians among extant terrestrial vertebrates, contributes to sharpening the peripheral matched filter by improving the signal-to-noise ratio for perceiving conspecific calls within a noisy multi-species breeding chorus (Fig. 2). Recordings with a laser Doppler vibrometer showed that inflated lungs had a peak resonance at a frequency value that fell in a frequency range (1400-2200 Hz) that was between the spectral peaks of conspecific calls and that coincided with a suspected region of frequency overlap between the AP and BP (Fig. 2a). Compared to a lung-deflated condition, lung inflation attenuated the response of the eardrum at this same intermediate frequency range by about 4 dB to 10 dB (Lee et al. 2021) (Fig. 2b). Importantly, changes in lung inflation did not significantly affect the sensitivity of the eardrum to the spectral peaks present in conspecific calls. An analysis of these lung-mediated effects with respect to a physiological model of the peripheral frequency tuning in green treefrogs (Lee et al. 2017b) suggested receivers might benefit from lung-mediated reductions in auditory masking by sound frequencies occurring between the two spectral peaks of the call (Fig. 2c-e). The question remained, however, as to how frequently green treefrog receivers might encounter frequencies in this critical range. A social network analysis of data from the North American Amphibian Monitoring Program (NAAMP) revealed that just 10 heterospecific species accounted for ~80% of the observed instances of "co-calling" between green treefrogs and another frog species. Of the 10 heterospecific species, five produce advertisement calls with spectral components that fall within the intermediate frequency range between the spectral peaks of *H. cinerea* calls. Thus, in the real-world environment of a mixed-species chorus, lung inflation could function to mitigate problems of auditory masking from a prominent subset of heterospecific calls by sharpening the tuning of the matched spectral filter in the periphery. That is, inflated lungs may function in enhancing the auditory contrast between conspecific signals and heterospecific noise. It will be important in future studies to test this hypothesis by measuring changes in the frequency tuning of different populations of auditory nerve fibers (e.g., Feng et al. 1975) that are expected to occur with changes in lung inflation.

Rising above the noise

Feng and his colleagues discovered and investigated a remarkable example of how environmental noise shapes the co-evolution of signalers and receivers, as expected according to the matched filter hypothesis, in this instance by selecting for a shift to using higher frequency signals (Feng et al. 2002; Narins et al. 2004; Suthers et al. 2006; Feng and Schul 2007; Feng and Narins 2008; Arch et al. 2008, 2009, 2011, 2012; Gridi-Papp et al. 2008). Concave-eared torrent frogs (*Odorrana tormota*, formerly *Amolops tormotus*) breed in habitats adjacent to fast-flowing streams and waterfalls. These habitats are characterized by high levels of broadband noise, with most acoustic energy around 100 Hz and extending to frequencies above 20 kHz (Narins et al. 2004). *O. tormata* males typically produce two types of calls: two-note calls consisting of two short (~100 ms) pips, and one-note calls of either short (~150 ms) or long (~400 ms) durations (Narins et al. 2004). Both types of calls are frequency modulated and often include sudden onsets and offsets of harmonic and subharmonic components (Narins et al. 2004). Interestingly, the calls of this species include a first formant below ~30 kHz, a second formant at ~60 kHz, and sometimes a third formant at 105 kHz (Narins et al. 2004). Importantly, behavioral and electrophysiological experiments have shown that males can hear the ultrasonic components of the calls and adjust their behavior in response to them (Feng et al. 2006). Compared to spontaneous calling activity, males increased their call rate in response to playbacks of only the audible (< 20 kHz) or only the ultrasonic (> 20 kHz) components of the calls (Feng et al. 2006) (Fig. 3a). Auditory evoked potentials and single unit recordings from the auditory midbrain revealed hearing sensitivity to sounds with frequencies up to 34 kHz (Feng et al. 2006) (Fig. 3b). Furthermore, gravid females also produce calls with ultrasonic elements, and males show remarkably acute phonotaxis in response to playbacks of female calls (Shen et al. 2008). These studies by Feng and colleagues provided the first record of ultrasonic communication in anurans, which has now also been demonstrated in the hole-in-the-head frog (*Huia cavitympanum*) an endemic Bornean frog that also breeds in habitats with high levels of noise (Arch et al. 2008, 2009). Together, these studies of ultrasonic signaling in frogs beautifully demonstrate adaptations that facilitate communication in noisy environments based on exploiting separation in frequency between signals and noise. Other articles in this special issue honoring Albert Feng provide more in-depth treatments examining the evolution of high-frequency calls (Xiong & Jiang this volume) and the mechanisms of high-frequency hearing (Cobo-Cuan et al. this volume). Previous reviews provide information on additional ways that frogs behaviorally respond to noise, for example, by altering their vocal behavior with respect to neighboring frogs in a chorus (Gerhardt and Huber 2002; Wells and Schwartz 2007; Schwartz and Bee 2013).

A limitation of matched spectral filtering

For many frog species, the main source of auditory masking is not the calls of other species or other sources of environmental noise, but the calling of *conspecific* males (Fig. 4). This is because matched spectral filtering is of limited use in coping with noise generated by conspecific signals that coevolved to pass through the filter. As noted above, problems of auditory masking are most severe when competing sounds have similar spectral content, occur at the same time, and originate from the same location. For any given frog species, the spectral content of vocalizations used for intraspecific communication will generally fall within a restricted frequency range that is audible to all adult members of the species. Moreover, mixed-species choruses often consist of different species whose calls may be well segregated in frequency (Wells 2007; Schwartz and Bee 2013). Although the males of some frog species shift the timing of their calls to avoid overlap with immediately neighboring conspecifics in the chorus (Schwartz and Bee 2013), this behavior is not characteristic of all species (e.g., Schwartz et al. 2002) and it does not lead to complete avoidance of call overlap with all neighbors, particularly in dense choruses (e.g., Greenfield and Rand 2000). While calling neighbors usually attempt to maintain some minimum distance from other conspecific males (e.g., Wilczynski and Brenowitz 1988), these distances shrink as chorus density increases, and there is evidence of species-specific clustering within mixed-species choruses (Wells 2007). In dense choruses of conspecifics, behaviorally important features of frog calls (e.g., a pulsed structure) can become obscured by the background noise of the chorus and overlapping calls produced by nearby males (Kuczynski et al. 2010). To cope with auditory masking resulting from the calls of conspecifics, frogs must employ mechanisms other than matched spectral filtering, which we discuss in the next two sections.

Exploiting temporal fluctuations in noise

A second major theme running through Feng's research on frogs was to discover neural mechanisms for processing complex, temporally-patterned sounds, particularly the species-specific temporal properties of conspecific vocalizations (Hall and Feng 1986, 1988; Condon et al. 1991; Feng et al. 1991; Gooler and Feng 1992; Penna et al. 1997, 2001). Later in his career, Feng noted the importance of understanding how chorus noise interferes with the temporal structure of anuran vocalizations and affects communication (Christie et al. 2010, 2019). An important extension of his work on the temporal processing of sounds and on communication in noisy environments was his investigation into how the auditory nervous system detects signals in temporally modulated noise (Goense and Feng 2012).

Real-world sounds are amplitude modulated and, often, comodulated. This means that their sound levels fluctuate over time and that amplitude fluctuations are often correlated across different regions of the frequency spectrum (Richards and Wiley 1980; Nelken et al. 1999). Decades of psychophysical studies provide overwhelming evidence that the human auditory system exploits amplitude modulations and comodulation in background noise to solve the cocktail party problem (Verhey et al. 2003). Surprisingly, these well-known features of natural sounds, and how the auditory system may exploit them, are often neglected in studies of animal communication in noise. Feng was among the first neuroethologists to draw attention to this problem and to investigate how the anuran auditory system may exploit amplitude fluctuations in background noise to facilitate signal detection and recognition in real-world environments (Feng and Schul 2007; Goense and Feng 2012).

Dip listening

Our ability to catch short "acoustic glimpses" of target signals when the amplitude of background noise momentarily drops is known as "dip listening" (Buus 1985; Cooke 2006; Vestergaard et al. 2011). When masker levels fluctuate in time, speech recognition thresholds are usually lower than those in the presence of non-fluctuating maskers (Gustafsson and Arlinger 1994; Bacon et al. 1998). Maskers with slow rates of amplitude modulation usually have longer dips in noise levels than those with faster rates, increasing the probability of catching meaningful glimpses of the target signals. Accordingly, masking release by means of dip listening is generally greater in the presence of maskers with slower, compared to faster, fluctuation rates (Gustafsson and Arlinger 1994; Bacon et al. 1998).

In breeding choruses of frogs and toads, the noise generated by the aggregation of calling males is amplitude modulated (Vélez and Bee 2010). Importantly, patterns of level fluctuations in chorus noise differ among species, reflecting species-specific properties of the mating call and calling behavior (Vélez and Bee 2010). A series of studies investigated dip listening in Cope's gray treefrogs (*Hyla chrysoscelis*) and green treefrogs (*H. cinerea*). In the presence of sinusoidally amplitude modulated (SAM) chorus-shaped maskers, green treefrogs did not benefit from dip listening; signal recognition thresholds were similar to those in the presence of a non-fluctuating masker (Vélez et al. 2012). In contrast, Cope's gray treefrogs experienced a 2-4 dB release from masking when SAM maskers fluctuated at slow rates (e.g., <5 Hz; Fig. 5) (Vélez and Bee 2011). Signal recognition thresholds were not different from those in the non-fluctuating control when SAM maskers fluctuated at intermediate rates (e.g., 5-20 Hz; Fig. 5), and 4-6 dB higher in the presence of maskers with high rates of fluctuation (e.g., 40-80 Hz; Fig. 5)(Vélez and Bee 2011). Dip listening can explain the release from masking observed in slowly fluctuating maskers. In quiet conditions, females respond to calls with at least six to nine pulses, but not to calls with five or fewer pulses (Vélez and Bee 2011; Gupta et al. 2021). At slow rates of fluctuation, dips in masker levels allowed for glimpses of nine or more consecutive pulses of the advertisement call; at intermediate and high rates of fluctuation, dips allowed for glimpses of five or fewer consecutive pulses (Vélez and Bee 2011). Furthermore, the additional masking experienced by females in the presence of SAM maskers with high rates of fluctuation (e.g., 40 Hz and 80 Hz; Fig. 5) is consistent with a phenomenon known as modulation masking (Bacon and Grantham 1989; Kwon and Turner 2001). The pulse rate of the call is approximately 40 to 60 pulses/s and an important acoustic property used by females for species recognition (Schul and Bush 2002). Hence, the temporal structure of SAM maskers at faster rates may have interfered with processing the temporal structure of the advertisement call (Schwartz and Marshall 2006; Marshall et al. 2006).

While Cope's gray treefrog may listen in the dips of slowly fluctuating SAM maskers, this ability may provide limited benefits for communicating in real-world situations. Vélez and Bee (2013) showed that Cope's gray treefrog and green treefrog females do not exploit natural amplitude fluctuations in chorus noise to recognize male mating calls. Furthermore, compared to non-fluctuating maskers, the ability of Cope's gray treefrog females to discriminate between calls varying in duration does not improve in the presence of maskers with natural or sinusoidal amplitude modulations (Vélez et al. 2013a). Interestingly, however, females of the closely-related eastern gray treefrog (*Hyla versicolor*) experienced a ~2-6 dB release from masking in the presence of naturally fluctuating maskers compared to non-fluctuating maskers during call recognition tests (Schwartz et al. 2013). This dip-listening effect was similar for target signals varying in call length (calls of 10, 20, 30, or 40 pulses) with equal or unequal pulse efforts (Schwartz et al. 2013). Additional evidence for dip listening has come from studies of comodulation masking release.

Comodulation masking release

Studies of comodulation masking release (CMR) reveal that human listeners experience a release from auditory masking when amplitude fluctuations in noise are correlated across the frequency spectrum, compared with conditions lacking fluctuations or when different frequency bands fluctuate independently (reviewed in Verhey et al. 2003). CMR was first attributed to a process by which the auditory system integrates energy across auditory filters to differentiate signals from noise (Hall et al. 1984); however, subsequent studies revealed that CMR also depends on within auditory-filter mechanisms (e.g., Schooneveldt and Moore 1987). In humans, the effect of CMR is usually larger when maskers have (i) large bandwidths, (ii) slow modulation rates, (iii) high modulation depths, (iv) irregular fluctuations, and (v) high levels (reviewed in Verhey et al. 2003).

Comodulation is likely a property of many natural sounds (Klump 1996; Nelken et al. 1999). In their review on hearing in real-world situations, Feng and Schul (2007) pointed out that

CMR should provide an advantage for listening in frog choruses. They suggested that chorus noise should be coherently modulated across the frequency spectrum given the repetitive nature of anuran calls and other biotic sounds, the tendency of males to avoid call overlap with their nearest neighbors, and the modulations imposed by wind. Indeed, Lee et al. (2017b) recently showed that the noise generated in Cope's gray tree frog choruses is comodulated. They developed an anatomical/physiological model of the auditory periphery, based on tuning curves of auditory nerve fibers, that simulated spectral processing by the amphibian and basilar papillae. When recordings of chorus noise were passed through this model, temporal modulations on the output of filters centered around 1.3 and 2.6kHz, which correspond to filters in the AP and BP, respectively, were highly correlated. These results show that temporal fluctuations in chorus noise are correlated across the frequency spectrum, and that the peripheral auditory system of frogs can potentially transduce these comodulations.

To date, only one study has investigated neural correlates of CMR in frogs. Goense and Feng (2012) tested the hypothesis that neurons in the auditory midbrain (inferior colliculus) of northern leopard frogs (*Rana pipiens*) contribute to CMR. Using a band-widening paradigm, the bandwidth of unmodulated and sinusoidally amplitude modulated (6.7 Hz modulation rate) maskers was systematically varied between 0.1 and 5 kHz. These masker bandwidths span the hearing range of northern leopard frogs (Mudry et al. 1977) and the range of auditory-filter bandwidths measured behaviorally for other species of frogs (Ehret and Gerhardt 1980; Narins 1982; Moss and Simmons 1986). The target signal was a train of pulses resembling the temporal structure of the conspecific advertisement call, centered at the unit's characteristic frequency. Overall, signal detection thresholds were lower in the presence of modulated maskers, compared to unmodulated maskers. In 10% of neurons that exhibit masking release, responses to the target signal were stronger during 'dips' of modulated maskers, providing evidence for neural correlates of dip-listening in frogs. Additionally, masking release was more pronounced at the narrowest and widest masker bandwidths. Importantly, masking release depended on both masker modulation and masker bandwidth for 15% of neurons, providing evidence for neural correlates of CMR in the anuran central auditory system. Goense and Feng (2012) suggested that such neurons may contribute to behavioral CMR in frogs.

Recent studies have shown that female Cope's gray treefrogs exploit comodulation in background noise to reduce communication errors in evolutionarily important contexts. In a series of psychophysical experiments, Lee et al. (2017b) tested the extent to which female frogs benefit from CMR in three different contexts: recognizing advertisement calls (Fig. 6a), discriminating between conspecific and heterospecific calls differing in pulse rate (Fig. 6b), and discriminating between conspecific calls differing in call effort, a sexually selected property (Fig. 6c). Female frogs were tested in the absence or presence of chorus-shaped maskers constructed by adding two narrow-band noises (400-Hz bandwidth) centered on the two spectral peaks present in advertisement calls (1.3 kHz and 2.6 kHz). In the unmodulated masking condition, no amplitude modulations beyond the inherent fluctuations present in narrow-band noise were imposed on the two narrow-band noises. The envelopes of low-pass filtered noises (12.5 Hz cutoff) were used to impose random amplitude fluctuations on the narrow-band noises and create two amplitude-modulated masking conditions: uncorrelated and comodulated. In the uncorrelated condition, each masker band was modulated with a different envelope, which yielded temporal envelopes that fluctuate independently across bands. In the comodulated condition, both masker bands were modulated with the same envelope, resulting in correlated amplitude fluctuations across frequency. Compared to the unmodulated condition, signal recognition thresholds were lower in both modulated conditions. Importantly, thresholds were 2.6 dB lower in the comodulated condition, compared to the uncorrelated condition (Fig. 6a). When given a choice between conspecific and heterospecific calls, females preferred the conspecific call more often than expected by chance in both modulated masking conditions, but not in the unmodulated masking condition (Fig. 6b). Importantly, the proportion of females choosing the conspecific call was significantly higher in the comodulated masking condition, compared to the uncorrelated masking condition. Similarly, the proportion of females choosing high- over low-effort calls was higher in the comodulated condition compared to the uncorrelated and unmodulated masking conditions (Fig. 6c). These results confirmed a role for comodulation masking release in call recognition and discrimination.

In a follow-up study, Bee and Vélez (2018) used the same three masking conditions to investigate the effects of masker level on CMR. Signal recognition thresholds were obtained for each of the three maskers broadcast at 53 or73 dB SPL, and used to compute signal-to-noise ratios at threshold. Overall, signal-to-noise ratios were lower in the comodulated masking condition, and at the high, 73-dB, masker level. These results corroborate those from Lee et al. (2017b) that female frogs experience CMR, and are in line with studies from human listeners showing stronger effects of CMR at higher masker levels (Moore and Shailer 1991; Bacon et al. 1997). Together, results from these studies suggest that frogs can exploit comodulations, a natural scene statistic of frog breeding choruses and other natural sounds, to mitigate the impacts of auditory masking in ecologically relevant tasks of call recognition and discrimination.

The ability to exploit spectro-temporal correlations in noise to improve signal detection and recognition may have evolved early in vertebrate hearing. In addition to frogs, CMR has been documented through behavioral experiments in goldfish (Fay 2011), European starlings (Klump 2016), mice (Klink et al. 2010), gerbils (Klump et al. 2001), and dolphins (Branstetter and Finneran 2008). Similarly, neural correlates of CMR have also been reported in European starlings (Klump 2016), mice (Sollini and Chadderton 2016), gerbils (Diepenbrock et al. 2017), guinea pigs (Pressnitzer et al. 2001), and cats (Nelken et al. 1999). Whether the underlying mechanisms of CMR are common across vertebrates remains to be determined. CMR likely depends on a combination of within- and across-channel mechanisms (reviewed in Verhey et al. 2003). The anuran peripheral auditory system, with two auditory papillae, offers a unique opportunity to investigate the relative contributions of within- and across-channel mechanisms involved in CMR. Future work should integrate behavioral and physiological studies to uncover the underlying mechanisms of CMR that contribute to hearing in real-world situations.

Exploiting spatial separation between signals and noise

A third major theme in Feng's research on frog hearing – beginning with his dissertation (Feng 1975) – aimed to discover the neurosensory mechanisms of sound localization and directional hearing more broadly (Feng and Capranica 1976; Feng et al. 1976; Feng 1980, 1981; Feng and Shofner 1981; Gooler et al. 1993; Xu et al. 1994, 1996; Ratnam and Feng 1998; Lin and Feng 2001, 2003). This body of research, which Gerhardt et al. (this volume) review in the context of Feng's work on sound localization in frogs, inspired investigations into how frogs exploit spatial separation between signals and noise as a means of coping with auditory masking.

In humans, signal detection and recognition are improved when signals and noise originate from different locations in space compared to conditions in which they come from the same location. This so-called "spatial release from masking" (reviewed in Litovsky 2012) plays important roles in our ability to follow conversations in crowded social environments

(Bronkhorst 2000). When sources of speech and speech-like noise are separated in azimuth by 90°, for example, listeners often experience improvements in speech recognition thresholds of 6 dB to 12 dB or more (Bronkhorst 2000). Spatial release from masking arises when spatial separation between signals and noise creates a monaural cue that improves the signal-to-noise ratio at one ear (the so-called "best ear for listening") as well as binaural disparities between signal and noise in terms of interaural time and interaural level differences at the two ears. Spatial attention also plays a role in spatial release from masking.

Schwartz and Gerhardt (1989), citing the work of Feng and Shofner (1981) on the directional responses of frog auditory nerve fibers, were the first to examine spatial release from masking in frogs. Using a phonotaxis assay, they showed that female green treefrogs (*H. cinerea*) experience about 3 dB of masking release when sources of signals (advertisement calls and aggressive calls) and sources of broadband noise were separated by 45° to 90°. Spatial separation, however, did not improve females' discrimination between advertisement and aggressive calls, a feat they readily perform in quiet. In a subsequent series of electrophysiological studies of the larger northern leopard frog (*R. pipiens*), Feng and colleagues investigated the neural correlates of spatial release from masking in auditory nerve fibers and in the auditory midbrain (Ratnam and Feng 1998; Lin and Feng 2001, 2003). Ratnam and Feng (1998), for example, first showed that the detection thresholds of some frog midbrain neurons decreased when sources of signals and noise were separated in azimuth (see also Schwartz and Gerhardt 1995). Lin and Feng (2001) subsequently showed that auditory nerve fibers exhibit less spatial release from masking (average maximum of 2.9 dB; Fig 7a) compared with midbrain neurons (average maximum of about 9.4 dB; Fig. 7b), although there was considerable overlap in the magnitudes of masking release between both types of cells (Fig. 7). This important result suggested the hypothesis that additional processing by the central nervous system played an important role in exploiting spatial separation between signals and noise. Lin and Feng (2003) confirmed this hypothesis by showing that blocking $\mathrm{GABA_A}$ receptors, thereby reducing inhibition, in the midbrain reduced the magnitude of spatial release from masking to a magnitude more typical of auditory nerve fibers. Their interpretation of this result was that abolishing binaural inhibition reduced the sensitivity of midbrain neurons to interaural level differences, and thus to sound direction, thereby reducing the magnitude of neural spatial release from masking (Ling and Feng 2003).

More recent behavioral studies of spatial release from masking have been conducted in females of Cope's gray treefrog (*H. chrysoscelis*) using a phonotaxis paradigm. Based on the latency of phonotaxis, Bee (2007) estimated the magnitude of spatial release from masking to be on the order of 6 dB to 12 dB when sources of advertisement calls and chorus-shaped noise were separated by 90° compared with a co-located condition (Fig. 8). In a follow-up study, Nityananda and Bee (2012) used an adaptive tracking procedure (Bee and Schwartz 2009) to measure behavioral recognition thresholds in co-located and 90° separated conditions. On average, females experienced about 4.5 dB of masking release in the separated condition. This magnitude of masking release is smaller than might have been expected based on neural recordings from the leopard frog midbrain (Lin and Feng 2001), but it coincided closely with the magnitude of masking release expected to result solely from the inherent directionality of the gray treefrog's tympanum (Caldwell et al. 2016) and that based on recordings of leopard frog auditory nerve fibers (Lin and Feng 2001). Additional studies have shown that the benefits of spatial separation between signals and noise extend beyond reduced thresholds for detection and recognition to include improvements in sound pattern recognition. For example, spatial separation between signals and noise improved females' ability to discriminate between conspecific calls and those of a closely-related sister species (*H. versicolor*) based on differences in pulse rate (Bee 2008; Ward et al. 2013).

Together, results from behavioral and neurophysiological studies of frogs provide robust support for the hypothesis (Fay and Feng 1987; Feng and Ratnam 2000; Feng and Schul 2007) that spatial release from masking facilitates vocal communication in the noisy, real-world environment of a frog breeding chorus. At present, however, we lack a coherent understanding of how mechanisms underlying spatial release from masking produce perceptual benefits for receivers. For example, to what extent is the spatial release from masking demonstrated in behavior created by the inherent directionality of the peripheral auditory system versus augmented by additional processing in the central nervous system. To answer such questions, additional work is needed that examines the neural basis of spatial release from masking in species of frogs where behavioral data are available. It will also be important in future studies to understand the extent to which mechanisms for exploiting spatial separation between signals and noise interact with those that functionally exploit frequency separation between signals and noise as well as temporal fluctuations in noise amplitude.

Looking forward

In this final section, we discuss a few emerging and future research areas, both conceptual and methodological, that would build on Feng's considerable legacy to better understand the mechanisms and evolution of hearing and sound communication in frogs, particularly in terms of how they cope with noise. The topics discussed here are by no means exhaustive, but represent our suggestions for how forward progress could be made that would significantly enhance knowledge of frogs and their importance as vertebrate research systems in auditory neuroethology.

Energetic masking versus informational masking

Historically, students of animal communication have tended to use "masking" or "auditory masking" or "masking interference" as convenient catchall terms to describe the negative impacts of noise (Brumm and Slabbekoorn 2005; Barber et al. 2010; Luther and Gentry 2013). Indeed, we have also done so up to this point. This is in contrast to studies of human hearing and speech perception in noise, which often distinguish between the separate effects of two broad types of masking involving different underlying mechanisms, *energetic masking* and *informational masking* (Kidd et al. 2008). Energetic masking typically refers to masking that occurs when signals and noise overlap in time and frequency and are processed by the same auditory filter. Under conditions of energetic masking, excitation patterns caused by the noise plus signal cannot be distinguished from those caused by the noise alone. Quite commonly, energetic masking is viewed as occurring in the peripheral auditory system, for example when noise alone and noise plus signal generate indistinguishable response patterns in auditory nerve fibers, but it can also occur in the central nervous system (Durlach et al. 2003). In contrast to energetic masking, informational masking (Kidd et al. 2008) refers to situations when signals are audible (i.e., detected) but noise interferes with central auditory mechanisms that contribute to processing informative features of the signal. Informational masking is frequently associated with limitations of selective attention (e.g., when noise "distracts" a listener) or breakdowns in auditory grouping (e.g., when signals and noise are perceptually grouped together). Energetic masking and informational masking represent important but mechanistically distinct contributors to the difficulty we humans have following a conversation in social environments where multiple people are talking simultaneously.

To achieve a deeper understanding of how animals are adapted to cope with noise, future research on nonhuman animals, including frogs, must distinguish between energetic and informational masking in terms of their proximate causes and ultimate consequences. Presently, few studies of hearing and sound communication in animals properly distinguish between energetic and informational masking, even though the latter may be more pronounced under some circumstances (Rosa and Koper 2018). The available evidence from animal behavior studies conducted outside the context of acoustic communication suggests nonhuman animals also encounter problems of informational masking (e.g., Allen et al. 2021). Some of the results discussed in a previous section hint that acoustic signaling in frogs is also susceptible to informational masking. For example, in the study of dip listening in Cope's gray treefrog by (Vélez and Bee 2011), signal recognition thresholds in the presence of SAM noise were elevated when the modulation rate of the chorus-shaped masker was close to the pulse rate of the target signal (Fig. 5). Similar effects were not observed using spectrally-matched SAM maskers having lower and higher modulation rates. This result is difficult to explain as a function of energetic masking in the periphery. Instead, an informational masking interpretation of this result is that modulations in the SAM noise interfered with the ability of subjects to process the signal's pulse rate, which is an important species recognition cue in this species (Schul and Bush 2002). More recently, Gupta and Bee (2022) directly tested the informational masking hypothesis in the same species. They found that a sequence of slow-rate, random-frequency pulses that were temporally interleaved with those of the target signal but presented in a remote frequency range impaired call recognition relative to quiet and a control condition with a frequency- and level-matched bandlimited noise lacking temporal structure. Importantly, target signals and maskers were delivered in the separate frequency ranges primarily transduced by AP and BP, thereby significantly reducing the potential for energetic masking. The interpretation was that the pulsed structure of the informational masker interfered with processing the pulsed structure of the target signal. Additional studies of informational masking in frogs that integrate behavioral and electrophysiological approaches will be needed to understand how real-world noises interfere with the processing of biologically informative features of communication sounds.

Multimodal signaling

Over the past two decades, following the seminal paper by Partan and Marler (1999), multimodal signaling has become a hot topic in the study of animal communication (Hebets and Papaj 2005; Partan and Marler 2005; Higham and Hebets 2013; Halfwerk et al. 2019). Multimodal signals have multiple components that are transduced by receivers using more than one sensory modality. A prominent functional hypothesis for the evolution of multimodal signals is that they are favored because different components can convey either multiple messages or redundant messages (Johnstone 1996; Hebets and Papaj 2005; Elias et al. 2006). Multimodal signals may be particularly advantageous when noise in one sensory channel degrades the quality of signal information, but similar information can still be conveyed by redundant signal components transduced by a different sensory modality (Partan 2013, 2017).

Beginning with the robotic playback studies of Narins et al. (2003, 2005) and the video playback study by Rosenthal et al. (2004), frogs have featured prominently in studies of multimodal signaling. This work suggests receivers sometimes make different behavioral decisions when they perform “cross-modal integration,” that is, when they integrate acoustic information in advertisement calls with visual cues provided by a synchronously inflating vocal sac (reviewed in Farris and Taylor (2016)). Thus, we might expect reliance on the redundant visual cue of a vocal sac (see Fig. 1a) to help frogs mitigate noise problems in a way analogous to lipreading in humans (Grant and Seitz 2000; Bernstein et al. 2004). So far, however, few studies have investigated whether chorus noise influences the extent to which frogs rely on vocal sacs as redundant cues or signals. Evidence from recent robotic playback studies of túngara frogs (*Physalaemus pustulosus*)(Taylor et al. 2021) and Cope’s gray treefrogs (*H. chrysoscelis*)(Li et al. 2022) suggests visual cues associated with a vocal sac have limited efficacy in soliciting female approaches toward calling males in noisy conditions. In the study by Li et al. (2022), for example, the presence of a robotic calling male with a dynamically inflating vocal sac did not alter responses to acoustic signals with temporal structures degraded by the simulated effects of noise or to those broadcast in the presence of chorus noise. Additional work in more species, and with more realistic robots (Narins and Feng 2007), is needed to determine the extent to which frogs rely on visual cues to overcome cocktail-party-like problems. Much of the previous work on multimodal signaling in the context of mate choice in nocturnal frogs has used potentially problematic methods that should be avoided in future studies (reviewed in Li et al. 2022).

Anthropogenic noise

Feng emphasized the need to understand sound processing in "real world" environments (Feng and Ratnam 2000; Feng and Schul 2007). Sadly, those environments are changing rapidly. Urban development and the expansion of transportation networks are causing dramatic increases in the levels of human-generated noise that can substantially alter the soundscapes in which animals communicate (Barber et al. 2010; Shannon et al. 2016). Accordingly, the last two decades have seen an explosion of studies on whether and how animals adjust their signals and signaling behaviors in response to anthropogenic noise (Slabbekoorn et al. 2018; Kunc and Schmidt 2021; Gomes et al. 2022), including in frogs (e.g., Sun and Narins 2005). Recordings of anthropogenic noise (e.g., traffic noise; Fig. 9) near wetland breeding sites suggest frogs could be susceptible to human impacts on soundscapes. Compared to other taxa, however, relatively few studies have investigated the potential consequences of anthropogenic noise on frogs (reviewed in Simmons and Narins 2018; Kunc and Schmidt 2021; Gomes et al. 2022; Zaffaroni-Caorsi et al. 2022). As noted by (Narins and Feng 2007), it will be important to assess how the limited plasticity in frog calls and calling behavior allow them to cope with problems of anthropogenic noise. Studies on frogs and toads show mixed results, with strong differences among species in the acoustic properties affected by anthropogenic noise and in the magnitudes of the effect. It is therefore important to continue exploring this issue comparatively, considering ecological and morphological characteristics that may correlate with changes in calls or soundscape properties. Moving forward, we emphasize the need for using appropriate methods when comparing the amplitude and frequency of calls in the presence and absence of noise (Brumm and Zollinger 2011; Brumm et al. 2017), as well as adequate experimental designs to avoid problems with pseudoreplication (Kroodsma 1989).

Building upon Feng's research on hearing in real-world situations, future studies should also focus on how the auditory system of receivers adjusts to increasing levels of anthropogenic noise. Despite decades of research on the effects of anthropogenic noise on animal communication, we have very limited knowledge about how receivers from urban populations may cope with this problem (Slabbekoorn et al. 2018; Derryberry and Luther 2021). In frogs, few studies have investigated how anthropogenic noise impairs signal recognition by females (Zaffaroni-Caorsi et al. 2022). Whether receivers from populations exposed to high levels of anthropogenic noise experience similar or less masking than those from more rural areas is still unknown. With an AP that acts as a bank of band-pass filters sensitive to relatively low frequencies, and a BP that acts as one broad auditory filter tuned to higher frequencies (reviewed in Smotherman and Narins 2000), the anuran ear offers a superb opportunity to understand how different auditory systems may respond to human-induced changes to the soundscape. Populations of frogs and toads in small pockets of suitable breeding habitat within cities, with potentially little gene flow between them and with other rural populations, set the stage for comparative studies on evolutionary responses, developmental plasticity, and short-term adjustments of the auditory system to novel soundscapes.

Awake-behaving preparations

As this review illustrates, numerous behavioral and neurophysiological studies have investigated the impacts of auditory masking in frogs. With a few notable exceptions (e.g., Simmons 1988; Freedman et al. 1988; Schwartz & Gerhardt 1995), however, most of these behavioral and neurophysiological studies have been conducted not only using different individuals as subjects, but more often using individuals of different species. Most behavioral studies of auditory masking have been conducted in hylid treefrogs (*H. chrysoscelis*, *H. versicolor*, *H. cinerea,* and *H. gratiosa*), whereas most neurophysiological studies of auditory masking have used ranid frogs as subjects (e.g., *R. pipiens, R. catesbeiana*). To our knowledge, no study of frogs has investigated auditory masking – or any other aspect of hearing – in awake, behaving animals in which neural recordings are made simultaneously from individuals as they perform various tasks involving sound detection, recognition, discrimination, and localization.

Elucidating mechanisms that allow frogs to communicate in real-world environments will require overcoming a number of key challenges to enable neural recordings from awake, behaving animals. While frogs have proven suitable subjects in behavioral studies using reflex modification to measure auditory thresholds and critical ratios, this method has not yet been widely adopted, and more traditional psychoustic studies based on classical or operant conditioning have proven challenging (Simmons and Moss 1995). Therefore, most masking studies of frogs rely on evoking the animal's natural behaviors, such as evoked calling by males (e.g., Narins 1982) or phonotaxis by females, as reviewed here. Using phonotaxis as a behavioral assay in animals famous for saltatory locomotion presents several technological challenges for making stable neural recordings. Phonotaxis studies frequently require subjects to hop toward playback speakers located a meter or more away from a starting point. Mounting hardware such as microdrives or headstages to the thin skulls of small, jumping frogs poses obvious technical challenges (but see Mohammed et al. 2013). Frog brains are rather pulpous compared with those of birds and mammals, making it difficult to record from single units in behaving animals for any length of time using metal electrodes fixed to the skull. In addition, because phonotaxis requires frogs to move within the sound field, its use complicates experimental studies of some key phenomena of interest, such as spatial release from masking, that depend on manipulating spatial relationships between signals and noise. We suggest one way to potentially overcome these challenges would be to integrate the use of flexible neural probes (e.g., Zhao et al. 2019; Pimenta et al. 2021) to record from animals that are restrained from making large, saltatory movements but are still able to freely exhibit measurable phonotaxis-related behavior (Márquez et al. 2008; Gupta et al. 2020). Technological advances along these lines would, for the first time in frogs, allow researchers to relate auditory-evoked neural responses to the acoustically-guided behavioral decisions of individuals. Using such methods, we believe the field can substantially build upon Albert Feng's legacy by investigating how the frog's brain listens to sounds not only in the real-world but also in real time.

Acknowledgements

We thank Peter Narins for inviting us to contribute this article in memory of Albert Feng, whose work on hearing and sound communication in frogs has inspired our careers and research programs.

Author contributions NL, AV, MAB: contributed equally to writing this manuscript.

Funding Much of the original research funding on hylid treefrogs reviewed here was funded by grants to MAB from the National Institute on Deafness and Other Communication Disorders (R03DC008396 and R01DC009582) and the National Science Foundation (IOS-0842759, IOS-1452831). Preparation of this article was supported in part by grants to NL (IOS-2144831) and MAB (IOS-2022253) from the National Science Foundation.

Declarations

Conflicts of interest The authors declare no competing or financial interests.

Ethical approval Ethical approval was not required for this article.

References

Allen LC, Hristov NI, Rubin JJ, et al (2021) Noise distracts foraging bats. Proc R Soc B Biol Sci 288:20202689. https://doi.org/10.1098/rspb.2020.2689

Arak A (1983) Sexual selection by male–male competition in natterjack toad choruses. Nature 306:261–262. https://doi.org/10.1038/306261a0

Arch V, Burmeister S, Feng A, et al (2011) Ultrasound-evoked immediate early gene expression in the brainstem of the Chinese torrent frog, *Odorrana tormota*. J Comp Physiol A Neuroethol Sens Neural Behav Physiol 197:667–675. https://doi.org/10.1007/s00359-011-0626-9

Arch V, Simmons D, Quinones P, et al (2012) Inner ear morphological correlates of ultrasonic hearing in frogs. Hear Res 283:70–79. https://doi.org/10.1016/j.heares.2011.11.006

Arch VS, Grafe TU, Gridi-Papp M, Narins PM (2009) Pure ultrasonic communication in an endemic Bornean frog. PLOS ONE 4:e5413. https://doi.org/10.1371/journal.pone.0005413

Arch VS, Grafe TU, Narins PM (2008) Ultrasonic signalling by a Bornean frog. Biol Lett 4:19–22. https://doi.org/10.1098/rsbl.2007.0494

Bacon SP, Grantham DW (1989) Modulation masking: Effects of modulation frequency, depth, and phase. J Acoust Soc Am 85:2575–2580. https://doi.org/10.1121/1.397751

Bacon SP, Lee J, Peterson DN, Rainey D (1997) Masking by modulated and unmodulated noise: Effects of bandwidth, modulation rate, signal frequency, and masker level. J Acoust Soc Am 101:1600–1610. https://doi.org/10.1121/1.418175

Bacon SP, Opie JM, Montoya DY (1998) The effects of hearing loss and noise masking on the masking release for speech in temporally complex backgrounds. J Speech Lang Hear Res 41:549–563. https://doi.org/10.1044/jslhr.4103.549

Barber JR, Crooks KR, Fristrup KM (2010) The costs of chronic noise exposure for terrestrial organisms. Trends Ecol Evol 25:180–189. https://doi.org/10.1016/j.tree.2009.08.002

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

Bee MA (2007) Sound source segregation in grey treefrogs: spatial release from masking by the sound of a chorus. Anim Behav 74:549–558. https://doi.org/10.1016/j.anbehav.2006.12.012

Bee MA (2008) Finding a mate at a cocktail party: spatial release from masking improves acoustic mate recognition in grey treefrogs. Anim Behav 75:1781–1791. https://doi.org/10.1016/j.anbehav.2007.10.032

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

Bee MA, Schwartz JJ (2009) Behavioral measures of signal recognition thresholds in frogs in the presence and absence of chorus-shaped noise. J Acoust Soc Am 126:2788–2801. https://doi.org/10.1121/1.3224707

Bee MA, Swanson EM (2007) Auditory masking of anuran advertisement calls by road traffic noise. Anim Behav 74:1765–1776. https://doi.org/10.1016/j.anbehav.2007.03.019

Bee MA, Vélez A (2018) Masking release in temporally fluctuating noise depends on comodulation and overall level in Cope’s gray treefrog. J Acoust Soc Am 144:2354–2362. https://doi.org/10.1121/1.5064362

Bernstein LE, Auer ET, Takayanagi S (2004) Auditory speech detection in noise enhanced by lipreading. Spec Issue Audio Vis Speech Process 44:5–18. https://doi.org/10.1016/j.specom.2004.10.011

Branstetter BK, Finneran JJ (2008) Comodulation masking release in bottlenose dolphins (*Tursiops truncatus*). J Acoust Soc Am 124:625–633

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

Bronkhorst AW (2000) The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions. Acustica 86:117–128

Brumm H (2013) Animal communication and noise. Springer, New York

Brumm H, Slabbekoorn H (2005) Acoustic communication in noise. In: Advances in the study of behavior. Academic Press, pp 151–209

Brumm H, Zollinger SA (2011) The evolution of the Lombard effect: 100 years of psychoacoustic research. Behaviour 148:1173–1198. https://doi.org/10.1163/000579511X605759

Brumm H, Zollinger SA, Niemelä PT, Sprau P (2017) Measurement artefacts lead to false positives in the study of birdsong in noise. Methods Ecol Evol 8:1617–1625. https://doi.org/10.1111/2041-210X.12766

Buerkle NP, Schrode KM, Bee MA (2014) Assessing stimulus and subject influences on auditory evoked potentials and their relation to peripheral physiology in green treefrogs (*Hyla cinerea*). Comp Biochem Physiol -Mol Integr Physiol 178:68–81. https://doi.org/10.1016/j.cbpa.2014.08.005

Buus S (1985) Release from masking caused by envelope fluctuations. J Acoust Soc Am 78:1958–1965. https://doi.org/10.1121/1.392652

Caldwell MS, Lee N, Bee MA (2016) Inherent directionality determines spatial release from masking at the tympanum in a vertebrate with internally coupled ears. J Assoc Res Otolaryngol 17:259–270. https://doi.org/10.1007/s10162-016-0568-6

Capranica RR, Moffat AJM (1983) Neurobehavioral correlates of sound communication in anurans. In: Ewert JP, Capranica RR, Ingle DJ (eds) Advances in Vertebrate Neuroethology. Springer US, Boston, MA, pp 701–730

Cherry EC (1953) Some experiments on the recognition of speech, with one and with two ears. J Acoust Soc Am 25:975–979. https://doi.org/10.1121/1.1907229

Christie K, Schul J, Feng A (2010) Phonotaxis to male’s calls embedded within a chorus by female gray treefrogs, *Hyla versicolor*. J Comp Physiol A Neuroethol Sens Neural Behav Physiol 196:569–579. https://doi.org/10.1007/s00359-010-0544-2

Christie K, Schul J, Feng A (2019) Differential effects of sound level and temporal structure of calls on phonotaxis by female gray treefrogs, *Hyla versicolor*. J Comp Physiol A Neuroethol Sens Neural Behav Physiol 205:223–238. https://doi.org/10.1007/s00359-019-01325-5

Condon CJ, Chang SH, Feng AS (1991) Processing of behaviorally relevant temporal parameters of acoustic stimuli by single neurons in the superior olivary nucleus of the leopard frog. J Comp Physiol A Neuroethol Sens Neural Behav Physiol 168:709–725

Cooke M (2006) A glimpsing model of speech perception in noise. J Acoust Soc Am 119:1562–1573. https://doi.org/10.1121/1.2166600

Derryberry EP, Luther D (2021) What is known—and not known—about acoustic communication in an urban soundscape. Integr Comp Biol 61:1783–1794. https://doi.org/10.1093/icb/icab131

Diepenbrock J-P, Jeschke M, Ohl FW, Verhey JL (2017) Comodulation masking release in the inferior colliculus by combined signal enhancement and masker reduction. J Neurophysiol 117:853–867. https://doi.org/10.1152/jn.00191.2016

Durlach NI, Mason CR, Kidd G, et al (2003) Note on informational masking (L). J Acoust Soc Am 113:2984–2987. https://doi.org/10.1121/1.1570435

Ehret G, Capranica RR (1980) Masking patterns and filter characteristics of auditory nerve fibers in the green treefrog (*Hyla cinerea*). J Comp Physiol A Neuroethol Sens Neural Behav Physiol 141:1–12. https://doi.org/10.1007/bf00611872

Ehret G, Gerhardt HC (1980) Auditory masking and effects of noise on responses of the green treefrog (*Hyla cinerea*) to synthetic mating calls. J Comp Physiol 141:13–18. https://doi.org/10.1007/bf00611873

Ehret G, Tautz J, Schmitz B, Narins PM (1990) Hearing through the lungs: Lung-eardrum transmission of sound in the frog *Eleutherodactylus coqui*. Naturwissenschaften 77:192–194. https://doi.org/10.1007/bf01131168

Elias DO, Lee N, Hebets EA, Mason AC (2006) Seismic signal production in a wolf spider: parallel versus serial multi-component signals. J Exp Biol 209:1074. https://doi.org/10.1242/jeb.02104

Farris HE, Taylor RC (2016) Mate searching animals as model systems for understanding perceptual grouping. In: Bee MA, Miller CT (eds) Psychological mechanisms in animal communication. Springer International Publishing, Cham, pp 89–118

Fay RR (2011) Signal-to-noise ratio for source determination and for a comodulated masker in goldfish, *Carassius auratus*. J Acoust Soc Am 129:3367–3372. https://doi.org/10.1121/1.3562179

Fay RR, Feng AS (1987) Mechanisms for directional hearing among nonmammalian vertebrates. In: Yost WA, Gourevitch G (eds) Directional hearing. Springer US, New York, NY, pp 179–213

Fay RR, Popper AN (2000) Evolution of hearing in vertebrates: the inner ears and processing. Hear Res 149:1–10. https://doi.org/10.1016/S0378-5955(00)00168-4

Feng A (1975) Sound localization in anurans : an electrophysiological and behavioral study. Ph.D. Dissertation, Cornell University

Feng A (1980) Directional characteristics of the acoustic receiver of the leopard frog (*Rana pipiens*) - A study of 8th nerve auditory responses. J Acoust Soc Am 68:1107–1114. https://doi.org/10.1121/1.384981

Feng A (1981) Directional response characteristics of single neurons in the torus semicircularis of the leopard frog (*Rana pipiens*). J Comp Physiol 144:419–428. https://doi.org/10.1007/BF00612574

Feng A, Capranica R (1976) Sound localization in anurans. 1. Evidence of binaural interaction in dorsal medullary nucleus of bullfrogs (*Rana catesbeiana*). J Neurophysiol 39:871–881. https://doi.org/10.1152/jn.1976.39.4.871

Feng A, Gerhardt H, Capranica R (1976) Sound localization behavior of green treefrog (*Hyla cinerea*) and barking treefrog (*Hyla gratiosa*). J Comp Physiol 107:241–252. https://doi.org/10.1007/BF00656735

Feng A, Narins P (2008) Ultrasonic communication in concave-eared torrent frogs (*Amolops tormotus*). J Comp Physiol A Neuroethol Sens Neural Behav Physiol 194:159–167. https://doi.org/10.1007/s00359-007-0267-1

Feng A, Narins P, Capranica R (1975) Three populations of primary auditory fibers in bullfrog

(*Rana catesbeiana*) - their peripheral origins and frequency sensitivities. J Comp Physiol 100:221–229. https://doi.org/10.1007/BF00614532

Feng A, Narins P, Xu C, et al (2006) Ultrasonic communication in frogs. Nature 440:333–336. https://doi.org/10.1038/nature04416

Feng A, Narins P, Xu C-H (2002) Vocal acrobatics in a Chinese frog, *Amolops tormotus*. Naturwissenschaften 89:352–356. https://doi.org/10.1007/s00114-002-0335-x

Feng A, Ratnam R (2000) Neural basis of hearing in real-world situations. Annu Rev Psychol 51:699–725. https://doi.org/10.1146/annurev.psych.51.1.699

Feng AS, Hall JC, Siddique S (1991) Coding of temporal parameters of complex sounds by frog auditory-nerve fibers. J Neurophysiol 65:424–445. https://doi.org/10.1152/jn.1991.65.3.424

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

Feng AS, Shofner WP (1981) Peripheral basis of sound localization in anurans. Acoustic properties of the frog's ear. Hear Res 5:201–216. https://doi.org/10.1016/0378-5955(81)90046-0

Fuzessery ZM, Feng AS (1982) Frequency selectivity in the anuran auditory midbrain - single unit responses to single and multiple tone stimulation. J Comp Physiol 146:471–484. https://doi.org/10.1007/BF00609443

Fuzessery ZM, Feng AS (1983a) Frequency-selectivity in the anuran medulla - excitatory and inhibitory tuning properties of single neurons in the dorsal medullary and superior olivary nuclei. J Comp Physiol 150:107–119. https://doi.org/10.1007/BF00605294

Fuzessery ZM, Feng AS (1983b) Mating call selectivity in the thalamus and midbrain of the leopard frog (*Ranna p. pipiens* - single and multiunit analyses. J Comp Physiol 150:333–344. https://doi.org/10.1007/BF00605023

Fuzessery ZM, Feng AS (1981) Frequency representation in the dorsal medullary nucleus of the leopard frog, *Rana p. pipiens*. J Comp Physiol 143:339–347. https://doi.org/10.1007/BF00611171

Gall MD, Wilczynski W (2015) Hearing conspecific vocal signals alters peripheral auditory sensitivity. Proc R Soc B Biol Sci 282:20150749. https://doi.org/10.1098/rspb.2015.0749

Gerhardt HC (1974) The significance of some spectral features in mating call recognition in the green treefrog (*Hyla Cinerea*). J Exp Biol 61:229–241

Gerhardt HC (1975) Sound pressure levels and radiation patterns of the vocalizations of some North American frogs and toads. J Comp Physiol 102:1–12. https://doi.org/10.1007/bf00657481

Gerhardt HC (1981) Mating call recognition in the green treefrog (*Hyla cinerea*): Importance of two frequency bands as a function of sound pressure level. J Comp Physiol 144:9–16. https://doi.org/10.1007/bf00612792

Gerhardt HC, Höbel G (2005) Mid-frequency suppression in the green treefrog (*Hyla cinerea*): mechanisms and implications for the evolution of acoustic communication. J Comp Physiol A 191:707–714. https://doi.org/10.1007/s00359-005-0626-8

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

Gerhardt HC, Schwartz JJ (2001) Auditory tuning and frequency preferences in anurans. In: Ryan MJ (ed) Anuran communication. Smithsonian Institution Press, Washington, D.C., pp 73–85

Goense J, Feng A (2005) Seasonal changes in frequency tuning and temporal processing in single neurons in the frog auditory midbrain. J Neurobiol 65:22–36. https://doi.org/10.1002/neu.20172

Goense J, Feng A (2012) Effects of noise bandwidth and amplitude modulation on masking in frog auditory midbrain neurons. PLOS ONE 7:. https://doi.org/10.1371/journal.pone.0031589

Gomes L, Solé M, Sousa-Lima RS, Baumgarten JE (2022) Influence of anthropogenic sounds on insect, anuran and bird acoustic signals: A meta-analysis. Front Ecol Evol 10:

Gooler DM, Condon CJ, Xu JH, Feng AS (1993) Sound direction influences the frequency-tuning characteristics of neurons in the frog inferior colliculus. J Neurophysiol 69:1018–1030. https://doi.org/10.1152/jn.1993.69.4.1018

Gooler DM, Feng AS (1992) Temporal coding in the frog auditory midbrain: the influence of duration and rise-fall time on the processing of complex amplitude-modulated stimuli. J Neurophysiol 67:1

Gooler DM, Xu JH, Feng AS (1996) Binaural inhibition is important in shaping the free-field frequency selectivity of single neurons in the inferior colliculus. J Neurophysiol 76:2580–2594

Grant KW, Seitz P-F (2000) The use of visible speech cues for improving auditory detection of spoken sentences. J Acoust Soc Am 108:1197–1208. https://doi.org/10.1121/1.1288668

Greenfield MD, Rand AS (2000) Frogs have rules: selective attention algorithms regulate chorusing in *Physalaemus pustulosus* (Leptodactylidae). Ethology 106:331–347. https://doi.org/10.1046/j.1439-0310.2000.00525.x

Gridi-Papp M, Feng AS, Shen J-X, et al (2008) Active control of ultrasonic hearing in frogs. Proc Natl Acad Sci. https://doi.org/10.1073/pnas.0802210105

Gupta S, Alluri RK, Rose GJ, Bee MA (2021) Neural basis of acoustic species recognition in a cryptic species complex. J Exp Biol 224:jeb243405. https://doi.org/10.1242/jeb.243405

Gupta S, Bee MA (2022) Informational masking can constrain acoustically guided mating behavior. bioRxiv 2022.03.31.486641. https://doi.org/10.1101/2022.03.31.486641

Gupta S, Marchetto PM, Bee MA (2020) Customizable recorder of animal kinesis (CRoAK): A multi-axis instrumented enclosure for measuring animal movements. HardwareX 8:e00116. https://doi.org/10.1016/j.ohx.2020.e00116

Gustafsson HÅ, Arlinger SD (1994) Masking of speech by amplitude - modulated noise. J Acoust Soc Am 95:518–529. https://doi.org/10.1121/1.408346

Halfwerk W, Lea AM, Guerra MA, et al (2016) Vocal responses to noise reveal the presence of the Lombard effect in a frog. Behav Ecol 27:669–676. https://doi.org/10.1093/beheco/arv204

Halfwerk W, Varkevisser J, Simon R, et al (2019) Toward testing for multimodal perception of mating signals. Front Ecol Evol 7:

Hall J, Feng A (1988) Influence of envelope rise time on neural responses in the auditory system of anurans. Hear Res 36:261–276. https://doi.org/10.1016/0378-5955(88)90067-6

Hall J, Feng AS (1986) Neural analysis of temporally patterned sounds in the frog’s thalamus - processing of pulse duration and pulse repetition rate. Neurosci Lett 63:215–220. https://doi.org/10.1016/0304-3940(86)90358-7

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

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

Higham JP, Hebets EA (2013) An introduction to multimodal communication. Behav Ecol Sociobiol 67:1381–1388. https://doi.org/10.1007/s00265-013-1590-x

Johnstone RA (1996) Multiple displays in animal communication: ‘backup signals’ and ‘multiple messages.’ Philos Trans R Soc Lond B Biol Sci 351:329–338. https://doi.org/10.1098/rstb.1996.0026

Jørgensen MB (1991) Comparative studies of the biophysics of directional hearing in anurans. J

Comp Physiol A Neuroethol Sens Neural Behav Physiol 169:591–598. https://doi.org/10.1007/bf00193548

Jørgensen MB, Gerhardt HC (1991) Directional hearing in the gray treefrog *Hyla versicolor* - eardrum vibrations and phonotaxis. J Comp Physiol -Sens Neural Behav Physiol 169:177–183. https://doi.org/10.1007/bf00215864

Kidd G, Mason CR, Richards VM, et al (2008) Informational masking. In: Yost WA, Popper AN, Fay RR (eds) Auditory perception of sound sources. Springer US, Boston, MA, pp 143–189

Klink KB, Dierker H, Beutelmann R, Klump GM (2010) Comodulation masking release determined in the mouse (*Mus musculus*) using a flanking-band paradigm. J Assoc Res Otolaryngol 11:79–88. https://doi.org/10.1007/s10162-009-0186-7

Klump GM (1996) Bird communication in the noisy world. In: Ecology and evolution of acoustic communication in birds. Cornell University Press, Itthaca, NY, pp 321–338

Klump GM (2016) Perceptual and neural mechanisms of auditory scene analysis in the European starling. In: Bee MA, Miller CT (eds) Psychological mechanisms in animal communication. Springer International Publishing, Cham, pp 57–88

Klump GM, Kittel MC, Wagner E (2001) Comodulation masking release in the Mongolian gerbil. In: Abstracts of the twenty-fourth midwinter research meeting of the association for research in otolaryngology. St. Petersburg, FL, p 84

Kroodsma DE (1989) Suggested experimental designs for song playbacks. Anim Behav 37:600–609. https://doi.org/10.1016/0003-3472(89)90039-0

Kuczynski MC, Velez A, Schwartz JJ, Bee MA (2010) Sound transmission and the recognition of temporally degraded sexual advertisement signals in Cope's gray treefrog (*Hyla chrysoscelis*). J Exp Biol 213:2840–2850. https://doi.org/10.1242/jeb.044628

Kunc HP, Schmidt R (2021) Species sensitivities to a global pollutant: A meta-analysis on acoustic signals in response to anthropogenic noise. Glob Change Biol 27:675–688. https://doi.org/10.1111/gcb.15428

Kwon BJ, Turner CW (2001) Consonant identification under maskers with sinusoidal modulation: Masking release or modulation interference? J Acoust Soc Am 110:1130–1140. https://doi.org/10.1121/1.1384909

Lee N, Christensen-Dalsgaard J, White LA, et al (2021) Lung mediated auditory contrast enhancement improves the Signal-to-noise ratio for communication in frogs. Curr Biol 31:1488-1498.e4. https://doi.org/10.1016/j.cub.2021.01.048

Lee N, Schrode KM, Bee MA (2017a) Nonlinear processing of a multicomponent communication signal by combination-sensitive neurons in the anuran inferior colliculus. J Comp Physiol -Neuroethol Sens Neural Behav Physiol 203:749–772. https://doi.org/10.1007/s00359-017-1195-3

Lee N, Ward JL, Velez A, et al (2017b) Frogs exploit statistical regularities in noisy acoustic scenes to solve cocktail-party-like problems. Curr Biol 27:743–750

Li H, Schrode KM, Bee MA (2022) Vocal sacs do not function in multimodal mate attraction under nocturnal illumination in Cope's grey treefrog. Anim Behav 189:127–146. https://doi.org/10.1016/j.anbehav.2022.05.003

Lin W-Y, Feng A (2001) Free-field unmasking response characteristics of frog auditory nerve fibers: comparison with the responses of midbrain auditory neurons. J Comp Physiol A Neuroethol Sens Neural Behav Physiol 187:699–712. https://doi.org/10.1007/s00359-001-0241-2

Lin WY, Feng AS (2003) GABA is involved in spatial unmasking in the frog auditory midbrain. J Neurosci 23:8143–8151

Litovsky RY (2012) Spatial release from masking. Acoust Today 8:18–25

Lombard RE, Straughan IR (1974) Functional aspects of anuran middle ear structures. J Exp Biol 61:71–93. https://doi.org/10.1242/jeb.61.1.71

Luther D, Gentry K (2013) Sources of background noise and their influence on vertebrate acoustic communication. Behaviour 150:1045–1068

Márquez R, Bosch J, Eekhout X (2008) Intensity of female preference quantified through playback setpoints: call frequency versus call rate in midwife toads. Anim Behav 75:159–166. https://doi.org/10.1016/j.anbehav.2007.05.003

Marshall VT, Schwartz JJ, Gerhardt HC (2006) Effects of heterospecific call overlap on the phonotactic behaviour of grey treefrogs. Anim Behav 72:449–459. https://doi.org/10.1016/j.anbehav.2006.02.001

McDermott JH (2009) The cocktail party problem. Curr Biol 19:R1024–R1027

Megela-Simmons A, Moss CF, Daniel KM (1985) Behavioral audiograms of the bullfrog (*Rana catesbeiana*) and the green tree frog (*Hyla cinerea*). J Acoust Soc Am 78:1236–1244. https://doi.org/10.1121/1.392892

Middlebrooks JC, Simon JZ, Popper AN, Fay RR (2017) The auditory system at the cocktail party. Springer, Berlin

Miranda JA, Wilczynski W (2009a) Female reproductive state influences the auditory midbrain response. J Comp Physiol A 195:341–349. https://doi.org/10.1007/s00359-008-0410-7

Miranda JA, Wilczynski W (2009b) Sex differences and androgen influences on midbrain auditory thresholds in the green treefrog, Hyla cinerea. Hear Res 252:79–88. https://doi.org/10.1016/j.heares.2009.04.004

Mohammed HS, Radwan NM, Walkowiak W, Elsayed AA (2013) A miniature microdrive for recording auditory evoked potentials from awake anurans. Gen Physiol Biophys 32:381–388. https://doi.org/10.4149/gpb_2013032

Moore BCJ (2013) An introduction to the psychology of hearing. Brill

Moore BCJ, Shailer MJ (1991) Comodulation masking release as a function of level. J Acoust Soc Am 90:829–835. https://doi.org/10.1121/1.401950

Moss C, Simmons A (1986) Frequency selectivity of hearing in the green treefrog, *Hyla cinerea*. J Comp Physiol A 159:257–266. https://doi.org/10.1007/bf00612308

Mudry KM, Constantine-Paton M, Capranica RR (1977) Auditory sensitivity of the diencephalon of the leopard frog *Rana p. pipiens*. J Comp Physiol 114:1–13. https://doi.org/10.1007/BF00656805

Narins PM (1982) Effects of masking noise on evoked calling in the Puerto Rican Coqui (Anura: Leptodactylidae). J Comp Physiol 147:439–446. https://doi.org/10.1007/BF00612008

Narins PM, Ehret G, Tautz J (1988) Accessory pathway for sound transfer in a neotropical frog. Proc Natl Acad Sci 85:1508–1512

Narins PM, Feng AS (2007) Hearing and sound communication in amphibians: prologue and prognostication. In: Narins PM, Feng AS, Fay RR, Popper AN (eds) Hearing and sound communication in amphibians. Springer New York, New York, NY, pp 1–11

Narins PM, Feng AS, Lin W, et al (2004) Old World frog and bird vocalizations contain prominent ultrasonic harmonics. J Acoust Soc Am 115:910–913. https://doi.org/10.1121/1.1636851

Narins PM, Grabul DS, Soma KK, et al (2005) Cross-modal integration in a dart-poison frog. Proc Natl Acad Sci 102:2425–2429. https://doi.org/10.1073/pnas.0406407102

Narins PM, Hödl W, Grabul DS (2003) Bimodal signal requisite for agonistic behavior in a dart-poison frog, *Epipedobates femoralis*. Proc Natl Acad Sci 100:577–580. https://doi.org/10.1073/pnas.0237165100

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

Nityananda V, Bee MA (2012) Spatial release from masking in a free-field source identification task by gray treefrogs. Hear Res 285:86–97. https://doi.org/10.1016/j.heares.2012.01.003

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

https://doi.org/10.1126/science.283.5406.1272

Partan SR (2013) Ten unanswered questions in multimodal communication. Behav Ecol Sociobiol 67:1523–1539. https://doi.org/10.1007/s00265-013-1565-y

Partan SR (2017) Multimodal shifts in noise: switching channels to communicate through rapid environmental change. Anim Behav 124:325–337. https://doi.org/10.1016/j.anbehav.2016.08.003

Partan SR, Marler P (2005) Issues in the classification of multimodal communication signals. Am Nat 166:231–245. https://doi.org/10.1086/431246

Patterson RD, Green DM (2012) Auditory masking. In: Carterette E (ed) Handbook of perception: Vol. IV Hearing. Elsevier, Amsterdam, the Netherlands, pp 337–361

Penna M, Capranica RR, Somers J (1992) Hormone-induced vocal behavior and midbrain auditory sensitivity in the green treefrog, *Hyla cinerea*. J Comp Physiol A 170:73–82. https://doi.org/10.1007/BF00190402

Penna M, Lin W, Feng A (1997) Temporal selectivity for complex signals by single neurons in the torus semicircularis of *Pleurodema thaul* (Amphibia: Leptodactylidae). J Comp Physiol A Neuroethol Sens Neural Behav Physiol 180:313–328. https://doi.org/10.1007/s003590050051

Penna M, Lin W, Feng A (2001) Temporal selectivity by single neurons in the torus semicircularis of *Batrachyla antartandica* (Amphibia : Leptodactylidae). J Comp Physiol A Neuroethol Sens Neural Behav Physiol 187:901–912. https://doi.org/10.1007/s00359-001-0263-9

Pimenta S, Rodrigues JA, Machado F, et al (2021) Double-layer flexible neural probe with closely spaced electrodes for high-density in vivo brain recordings. Front Neurosci 15:

Pressnitzer D, Meddis R, Delahaye R, Winter IM (2001) Physiological correlates of comodulation masking release in the mammalian ventral cochlear nucleus. J Neurosci 21:6377–6386

Ratnam R, Feng AS (1998) Detection of auditory signals by frog inferior collicular neurons in the presence of spatially separated noise. J Neurophysiol 80:2848–2859

Richards DG, Wiley RH (1980) Reverberations and amplitude fluctuations in the propagation of sound in a forest: implications for animal communication. Am Nat 115:381–399. https://doi.org/10.1086/283568

Rosa P, Koper N (2018) Integrating multiple disciplines to understand effects of anthropogenic noise on animal communication. Ecosphere 9:e02127. https://doi.org/10.1002/ecs2.2127

Rosenthal GG, Rand AS, Ryan MJ (2004) The vocal sac as a visual cue in anuran communication: an experimental analysis using video playback. Anim Behav 68:55–58. https://doi.org/10.1016/j.anbehav.2003.07.013

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

Schul J, Bush SL (2002) Non-parallel coevolution of sender and receiver in the acoustic communication system of treefrogs. Proc R Soc Lond B Biol Sci 269:1847–1852. https://doi.org/10.1098/rspb.2002.2092

Schwartz J, Buchanan B, Gerhardt HG (2002) Acoustic interactions among male gray treefrogs, *Hyla versicolor* , in a chorus setting. Behav Ecol Sociobiol 53:9–19. https://doi.org/10.1007/s00265-002-0542-7

Schwartz JJ, Bee MA (2013) Anuran acoustic signal production in noisy environments. In: Brumm H (ed) Animal communication and noise. Springer, Berlin, Heidelberg, pp 91–132

Schwartz JJ, Crimarco NC, Bregman Y, Umeoji KR (2013) An investigation of the functional significance of responses of the gray treefrog (*Hyla versicolor*) to chorus noise. J

Herpetol 47:354–360. https://doi.org/10.1670/12-027

Schwartz JJ, Gerhardt HC (1989) Spatially mediated release from auditory masking in an anuran amphibian. J Comp Physiol -Sens Neural Behav Physiol 166:37–41

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

Schwartz JJ, Marshall VT (2006) Forms of call overlap and their impact on advertisement call attractiveness to females of the gray treefrog, *Hyla versicolor* . Bioacoustics 16:39–56. https://doi.org/10.1080/09524622.2006.9753563

Shannon G, McKenna MF, Angeloni LM, et al (2016) A synthesis of two decades of research documenting the effects of noise on wildlife. Biol Rev 91:982–1005. https://doi.org/10.1111/brv.12207

Shen J-X, Feng AS, Xu Z-M, et al (2008) Ultrasonic frogs show hyperacute phonotaxis to female courtship calls. Nature 453:914–916. https://doi.org/10.1038/nature06719

Simmons A (2013) “To Ear is Human, to Frogive is Divine”: Bob Capranica’s legacy to auditory neuroethology. J Comp Physiol A 199:169–182. https://doi.org/10.1007/s00359-012-0786-2

Simmons AM, Moss CF (1995) Reflex modification: a tool for assessing basic auditory function in anuran amphibians. In: Klump GM, Dooling RJ, Fay RR, Stebbins WC (eds) Methods in comparative psychoacoustics. Birkhäuser Basel, Basel, pp 197–208

Simmons AM, Narins PM (2018) Effects of anthropogenic noise on amphibians and reptiles. In: Slabbekoorn H, Dooling RJ, Popper AN, Fay RR (eds) Effects of anthropogenic noise on animals. Springer New York, New York, NY, pp 179–208

Slabbekoorn HW, Dooling RJ, Popper AN, Fay RR (2018) Effects of anthropogenic noise on animals. Springer, Berlin

Smotherman MS, Narins PM (2000) Hair cells, hearing and hopping: a field guide to hair cell physiology in the frog. J Exp Biol 203:2237–2246

Sollini J, Chadderton P (2016) Comodulation enhances signal detection via priming of auditory cortical circuits. J Neurosci 36:12299. https://doi.org/10.1523/JNEUROSCI.0656-16.2016

Sun JWC, Narins PM (2005) Anthropogenic sounds differentially affect amphibian call rate. Biol Conserv 121:419–427. https://doi.org/10.1016/j.biocon.2004.05.017

Suthers R, Narins P, Lin W, et al (2006) Voices of the dead: complex nonlinear vocal signals from the larynx of an ultrasonic frog. J Exp Biol 209:4984–4993. https://doi.org/10.1242/jeb.02594

Tanner JC, Bee MA (2019) Within-individual variation in sexual displays: signal or noise? Behav Ecol 30:80–91. https://doi.org/10.1093/beheco/ary165

Taylor RC, Wilhite KO, Ludovici RJ, et al (2021) Complex sensory environments alter mate choice outcomes. J Exp Biol 224:jeb233288. https://doi.org/10.1242/jeb.233288

Vélez A, Bee MA (2010) Signal recognition by frogs in the presence of temporally fluctuating chorus-shaped noise. Behav Ecol Sociobiol 64:1695–1709. https://doi.org/10.1007/s00265-010-0983-3

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

Vélez A, Bee MA (2013) Signal recognition by green treefrogs (*Hyla cinerea*) and cope’s gray treefrogs (*Hyla chrysoscelis*) in naturally fluctuating noise. J Comp Psychol 127:166–178. https://doi.org/10.1037/a0030185

Vélez A, Höbel G, Gordon NM, Bee MA (2012) Dip listening or modulation masking? Call recognition by green treefrogs (*Hyla cinerea*) in temporally fluctuating noise. J Comp Physiol A 198:891–904. https://doi.org/10.1007/s00359-012-0760-z

Vélez A, Linehan-Skillings BJ, Gu Y, et al (2013a) Pulse-number discrimination by Cope's gray treefrog (*Hyla chrysoscelis*) in modulated and unmodulated noise. J Acoust Soc Am 134:3079–3089. https://doi.org/10.1121/1.4820883

Vélez A, Schwartz JJ, Bee MA (2013b) Anuran acoustic signal perception in noisy environments. In: Brumm H (ed) Animal communication and noise. Springer, New York, pp 133–185

Verhey JL, Pressnitzer D, Winter IM (2003) The psychophysics and physiology of comodulation masking release. Exp Brain Res 153:405–417

Vestergaard MD, Fyson NRC, Patterson RD (2011) The mutual roles of temporal glimpsing and vocal characteristics in cocktail-party listening. J Acoust Soc Am 130:429–439. https://doi.org/10.1121/1.3596462

Ward JL, Buerkle NP, Bee MA (2013) Spatial release from masking improves sound pattern discrimination along a biologically relevant pulse-rate continuum in gray treefrogs. Hear Res 306:63–75. https://doi.org/10.1016/j.heares.2013.09.006

Wells KD (2007) The ecology and behavior of amphibians. University of Chicago Press, Chicago

Wells KD, Schwartz JJ (2007) The Behavioral Ecology of Anuran Communication. In: Narins PM, Feng AS, Fay RR, Popper AN (eds) Hearing and Sound Communication in Amphibians. Springer New York, New York, NY, pp 44–86

Wilczynski W, Brenowitz EA (1988) Acoustic cues mediate inter-male spacing in a neotropical frog. Anim Behav 36:1054–1063. https://doi.org/10.1016/S0003-3472(88)80065-4

Wiley R (2015) Noise Matters: The Evolution of Communication. Harvard University Press, Cambridge, MA

Xu J, Gooler DM, Feng AS (1996) Effects of sound direction on the processing of amplitude modulated signals in the frog inferior colliculus. J Comp Physiol -Sens Neural Behav Physiol 178:435–445

Xu JH, Gooler DM, Feng AS (1994) Single neurons in the frog inferior colliculus exhibit direction-dependent frequency-selectivity to isointensity tone bursts. J Acoust Soc Am 95:2160–2170. https://doi.org/10.1121/1.408677

Yost WA (2008) Perceiving sound sources. In: Yost WA, Popper AN, Fay RR (eds) Auditory perception of sound sources. Springer US, Boston, MA, pp 1–12

Zaffaroni-Caorsi V, Both C, Márquez R, et al (2022) Effects of anthropogenic noise on anuran amphibians. Bioacoustics 1–31. https://doi.org/10.1080/09524622.2022.2070543

Zhang H, Feng AS (1998) Sound direction modifies the inhibitory as well as the excitatory frequency tuning characteristics of single neurons in the frog torus semicircularis (inferior colliculus). J Comp Physiol -Sens Neural Behav Physiol 182:725–735. https://doi.org/10.1007/s003590050217

Zhang H, Xu J, Feng A (1999) Effects of GABA mediated inhibition on direction-dependent frequency tuning in the frog inferior colliculus. J Comp Physiol A Neuroethol Sens Neural Behav Physiol 184:85–98. https://doi.org/10.1007/s003590050308

Zhao Z, Li X, He F, et al (2019) Parallel, minimally-invasive implantation of ultra-flexible neural electrode arrays. J Neural Eng 16:035001. https://doi.org/10.1088/1741-2552/ab05b6

Figure legends

Fig. 1 Matched filtering in the green treefrog *Hyla cinerea*. **a** Male *H. cinerea* producing an advertisement call. **b** Top - oscillogram of an exemplar single-note call, bottom - spectrogram of the same single-note exemplar demonstrating a bimodal frequency spectrum. **c** Behavioral responses of receptive female *H. cinerea* to bimodal and unimodal calls in two-alternative choice experiments. Bars depict the proportion (±95% exact binomial confidence intervals) of females that chose a bimodal call over a unimodal alternative. **d** A comparison of behavioral (red) and neurophysiological midbrain (blue) audiogram with the frequency spectrum of the *H. cinerea* advertisement call (black). The behavioral audiogram was computed by averaging data from two audiograms determined by reflex modification published in Megela-Simmons et al. (1985). The midbrain audiogram was computed by averaging midbrain audiograms from Lombard and Straughan (1974), Miranda and Wilczynski (2009), and Penna et al. (1992). To compensate for different stimulus frequencies used across studies, linear interpolation was used to interpolate between sampled frequency values. **e** Exemplar single-unit recordings obtained from a combination-sensitive facilitation neuron in the inferior colliculus. A synthetic single-note call stimulus is plotted to indicate stimulus timing. This combination-sensitive exemplar responds with a greater number of spikes when presented with a bimodal call compared to a unimodal call. **f** Bars depict the median (±IQR) number of spikes/stimulus expressed as a percentage of the maximum stimulus-driven response. Combination-sensitive facilitation units exhibit a heightened level of activity that is greater than the linear summation of activity in response to both unimodal stimuli presented in isolation. **c**,**e**, and **f** redrawn from Lee et al. (2017a).

Fig. 2 Lung mediated sharpening of the matched filter reduces masking from heterospecific calls. **a** A comparison of lung resonance determined from laser Doppler vibrometry (thick blue line and shaded blue area, mean ± 95% CI), and frequency spectrum of the *H. cinerea* advertisement call (thick black line and shaded gray area, mean ± 1 SD). Peak lung resonance occurs in between the spectral peaks of the *H. cinerea* advertisement call. **b** Heatmaps depicting the mean vibration amplitudes of the right eardrum in response to free-field acoustic stimulation in the deflated (left) and inflated (middle) states of lung inflation across frequency and sound incidence angle. Subtracting the inflated state heatmap from the deflated state heatmap results in the difference heatmap (right), which represents the effect of lung inflation on the mechanical response of the eardrum. The black contour in the difference heatmap encloses frequencies and angles where attenuation of the eardrum's response equaled or exceeded -4 dB when the lungs were inflated compared with deflated. The dashed lines indicate the minimum (1400 Hz) and maximum (2200 Hz) frequencies enclosed by the contour. **c** Physiological model of the peripheral frequency tuning in *H. cinerea* (colors) compared with the *H. cinerea* advertisement call spectrum (black), and the frequency region of the lung-mediated reduction in eardrum sensitivity (shaded gray area). Tuning curves are depicted separately for suppressible low-frequency and non-suppressible mid-frequency fibers innervating the amphibian papilla (AP) and for high-frequency fibers innervating the basilar papilla (BP). Neural responses of low-frequency fibers can be suppressed by frequencies in the range of mid-frequency AP fibers. This suppression would reduce the ability of low-frequency fibers to encode the lower spectral peak of *H. cinerea* advertisement calls. Lung-mediated reduction in the eardrum's response to this mid-frequency region is expected to reduce the activity of mid-frequency non-suppressible units, thus sharpening the matched filter, and reducing the suppression of low-frequency fibers. **d** and **e** Top - depicts the frequency spectra of advertisement calls produced by *R. clamitans* and *H. gratiosa*, respectively, both of which are heterospecific species that co-occur with *H. cinerea*. Bottom - depicts the magnitude of reduction in dB of spectral peaks in these heterospecific advertisement calls. Thick black lines depict the mean spectra, and thin gray lines depict the spectra of individual calls. Dotted line and shaded area depict the frequency region of the lung-mediated reduction in eardrum response. Polar plots depict the mean ± 95% CI attenuation of the call spectral peak within the shaded gray area as a function of sound incidence angle. Redrawn from Lee et al. (2021).

Fig. 3 Ultrasonic hearing in concave-eared torrent frogs. **a** Median number of calls produced in by eight male subjects during a 3 minute no-stimulation period and in response to 3 minute playbacks of a conspecific vocalization that was spectrally filtered to include only the audible frequencies (< 20 kHz) or only the ultrasonic frequencies (> 20 kHz). **b** Neural recordings from two representative neurons in the auditory midbrain (inferior colliculus) in response to tone-burst. The tonic unit responded to 10–27 kHz and had a best frequency of 20 kHz. The phasic unit responded to 5–30 kHz and had a best frequency of 10 kHz. Redrawn from Feng et al.

(2006).

Fig. 4 Calling conspecifics represent a potent source of auditory masking in chorusing frogs. **a** Spectrogram showing a 10-s long recording made from the edge of a pond that contained a dense, active chorus of Cope's gray treefrog. The nearly continuous bands of noise centered on the two spectral peaks of the species' advertisement call (approximately 1.3 and 2.6 kHz) depict the noise of the chorus. **b** A 1-s long recording of a single advertisement call made at close range (~ 1 m) with a highly directional, shotgun microphone from a spatially isolated male calling on a different night when chorus activity was low. Note the clear pulsatile structure of the single call. The single call is shown as an inset for illustrative purposes only.

Fig. 5 Dip listening in Cope's gray treefrog. Bars depict the means ± 95% confidence intervals for the threshold differences in the presence of sinusoidally amplitude modulated (SAM) chorus-shaped noise relative to a non-fluctuating condition. The horizontal dashed line represents no difference (i.e. 0 dB) from the non-fluctuating condition. Values lower than 0 dB represent less masking than in the non-fluctuating condition, and values greater than 0 dB represent more masking. Thresholds were significantly lower in slowly modulated SAM noises (≤ 2.5 Hz) and significantly higher in fast modulated SAM noises with modulation rates close to the species-specific pulse rate of advertisement calls. *$p < 0.05$. Redrawn from Vélez and Bee (2011).

Fig. 6 Comodulation masking release in Cope's gray treefrog. Shown here is a comparison of responses to signals presented in quiet and in the presence of chorus-shaped noise that had either temporally correlated (comodulated) envelope fluctuations, uncorrelated envelope fluctuations, or no envelope fluctuations (unmodulated). **a** Bars depict the mean ± s.e.m. signal recognition thresholds determined using an adaptive tracking procedure. The horizontal dashed line indicates the level of performance relative to the condition with the highest threshold. **b** Bars depict the proportion (±95% exact binomial confidence intervals) of subjects choosing stimuli with conspecific pulse rates, P(conspecific pulse rate). Horizontal dashed line depicts the level of performance expected by chance (0.5) in a two-alternative choice test. **c** Bars depict the proportions (±95% exact binomial confidence intervals) of subjects choosing stimuli with relatively higher calling efforts, P(higher calling effort). Horizontal dashed line depicts the level of performance expected by chance (0.5) in a two-alternative choice test. Subjects generally performed better in comodulated noise compared with uncorrelated and unmodulated noise. Redrawn from Lee et al. (2017b).

Fig. 7 Spatial release from masking in the ascending auditory system of the northern leopard frog. Histograms show the distributions of the amounts of maximum spatial release from masking resulting from angular separation of signal and noise for **a** auditory nerve fibers and **b** neurons from the auditory midbrain (inferior colliculus). The means (shown by arrows) and standard deviations are shown above each graph. The magnitude of neural spatial release from masking was greater in the midbrain compared with the auditory nerve. Data from Lin and Feng (2001). Redrawn from Feng and Schul (2007).

Fig. 8 Spatial release from masking in Cope's gray treefrog. Bars depict the mean ± s.e.m. normalized latency to respond to a synthetic advertisement call in the presence of artificial chorus-shaped noise (i.e., noise with a frequency spectrum similar to a natural chorus) as a function of the signal-to-noise ratio. Signal and noise were presented either from adjacent speakers (co-located condition; gray bars) or from speakers separated by 90° (separated, white bars). Latencies were normalized to latencies in response to the same call presented in quiet at an amplitude equivalent to that of the +6 dB condition. Responses were significantly faster in the separated condition at signal-to-noise ratios of -6 dB and 0 dB. *$p < 0.05$. Redrawn from Bee (2007).

Fig. 9 Traffic noise overlaps frog calls in frequency and time. Depicted here are the sound pressure levels of traffic noise recorded at a wetland in the area of Minneapolis/St. Paul, MN, USA as a function of **a** time of night and **b** frequency. Recordings were made near the water's edge and from the approximate height of the water's surface every 30 min between 1800 h and 0000 h, which spans the time of evening and night when many frog choruses form. The recording site was located 15 m from a nearby highway, which was elevated about 5 m above the surface of the water. Sound pressure levels (LCF_{peak}, LCF_{max}, LC_{eq}, or LCF_{min}) were measured in ⅓-octave bands. Redrawn from Bee and Swanson Bee and Swanson (2007).

Fig. 1

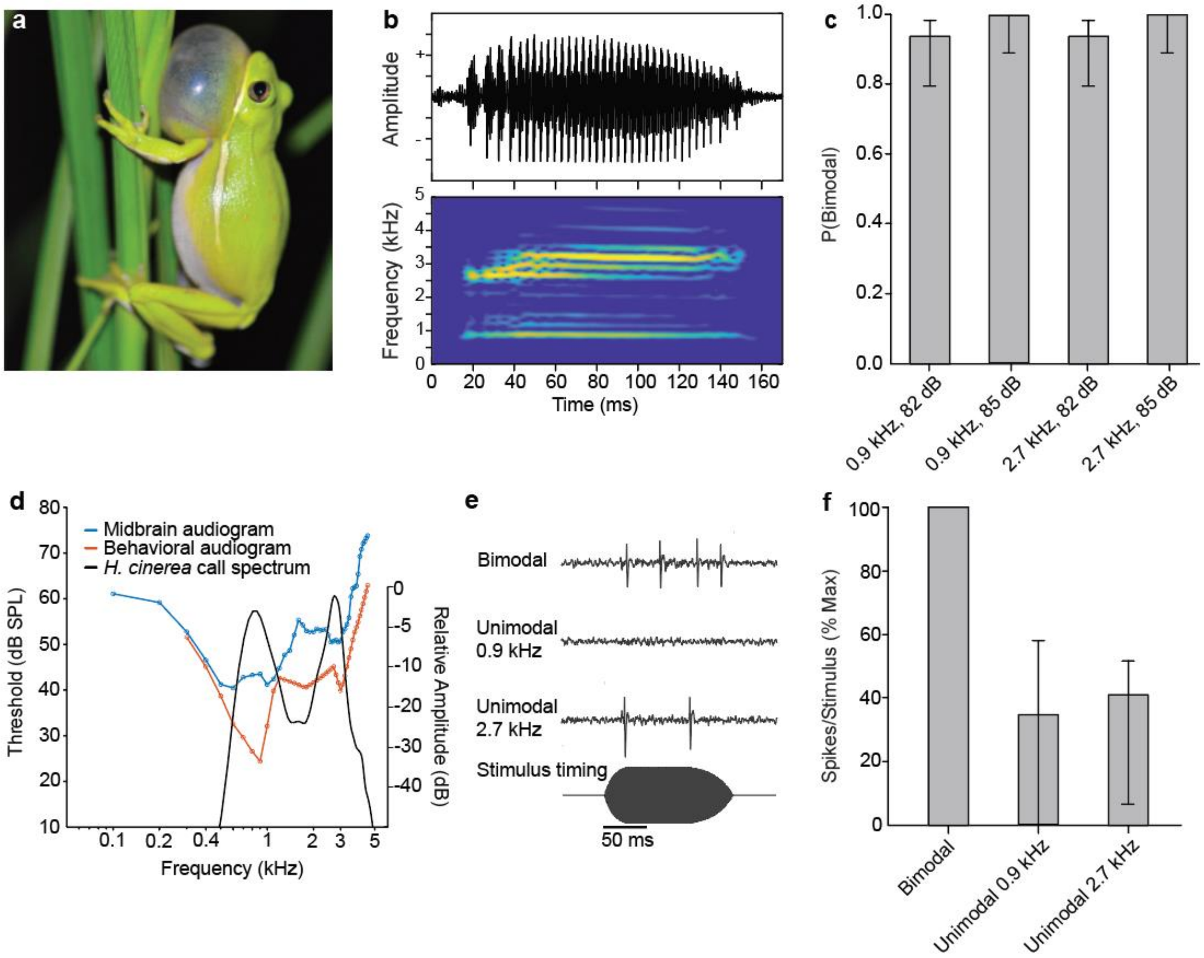

a
b
Amplitude
+
-
Frequency (kHz)
5
4
3
2
1
0
0 20 40 60 80 100 120 140 160
Time (ms)
c
P(Bimodal)
1.0
0.8
0.6
0.4
0.2
0.0
0.9 kHz, 82 dB
0.9 kHz, 85 dB
2.7 kHz, 82 dB
2.7 kHz, 85 dB
d
Midbrain audiogram
Behavioral audiogram
H. cinerea call spectrum
Threshold (dB SPL)
80
70
60
50
40
30
20
10
0.1 0.2 0.4 1 2 3 5
Frequency (kHz)
Relative Amplitude (dB)
0
-5
-10
-20
-30
-40
e
Bimodal
Unimodal 0.9 kHz
Unimodal 2.7 kHz
Stimulus timing
50 ms
f
Spikes/Stimulus (% Max)
100
80
60
40
20
0
Bimodal
Unimodal 0.9 kHz
Unimodal 2.7 kHz

Fig 2

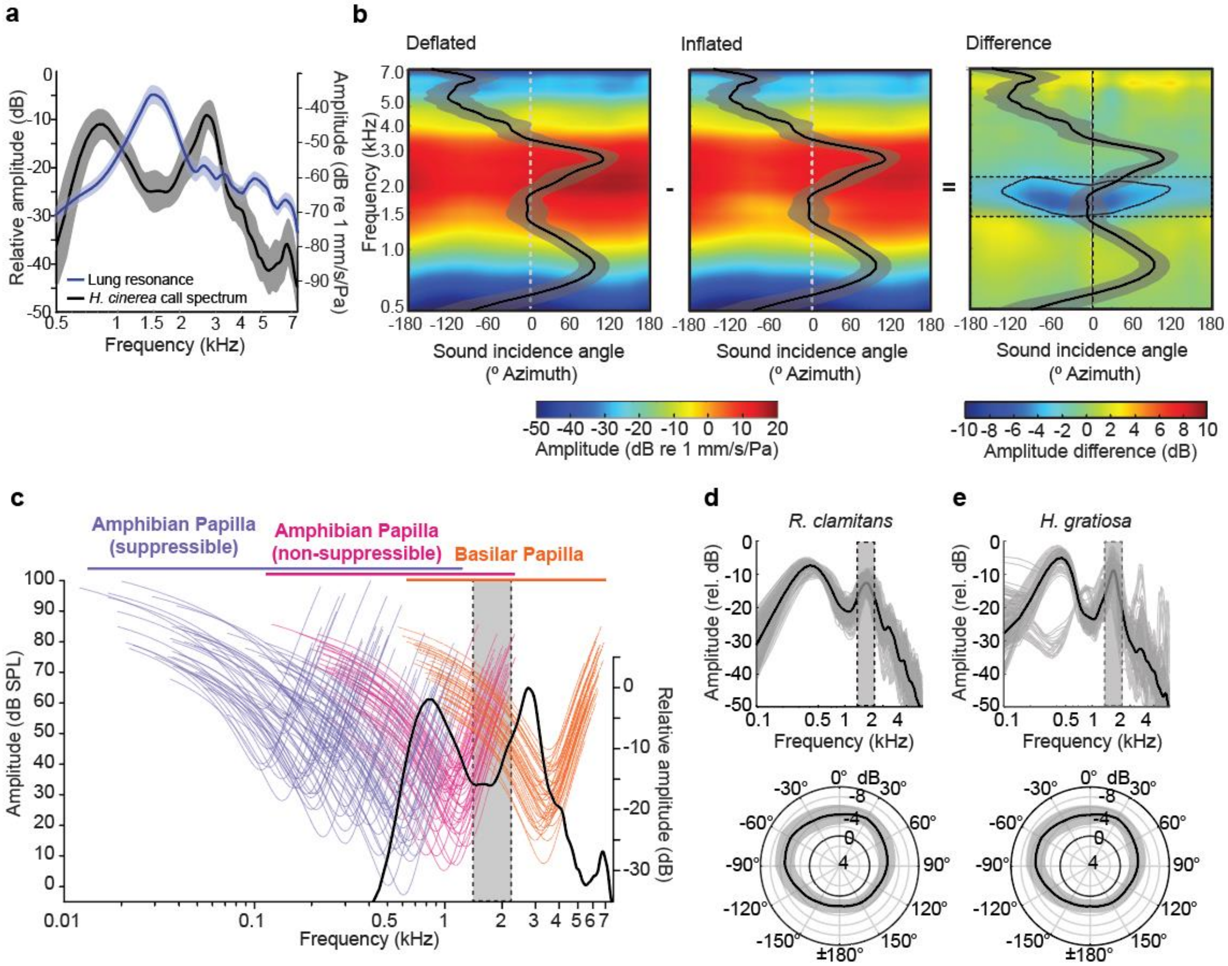

a
Relative amplitude (dB)
Amplitude (dB re 1 mm/s/Pa)
Lung resonance
H. cinerea call spectrum
Frequency (kHz)
b
Deflated
Inflated
Difference
Frequency (kHz)
Sound incidence angle (º Azimuth)
Amplitude (dB re 1 mm/s/Pa)
Amplitude difference (dB)
c
Amphibian Papilla (suppressible)
Amphibian Papilla (non-suppressible)
Basilar Papilla
Amplitude (dB SPL)
Relative amplitude (dB)
Frequency (kHz)
d
R. clamitans
Amplitude (rel. dB)
Frequency (kHz)
e
H. gratiosa
Amplitude (rel. dB)
Frequency (kHz)

Fig. 3

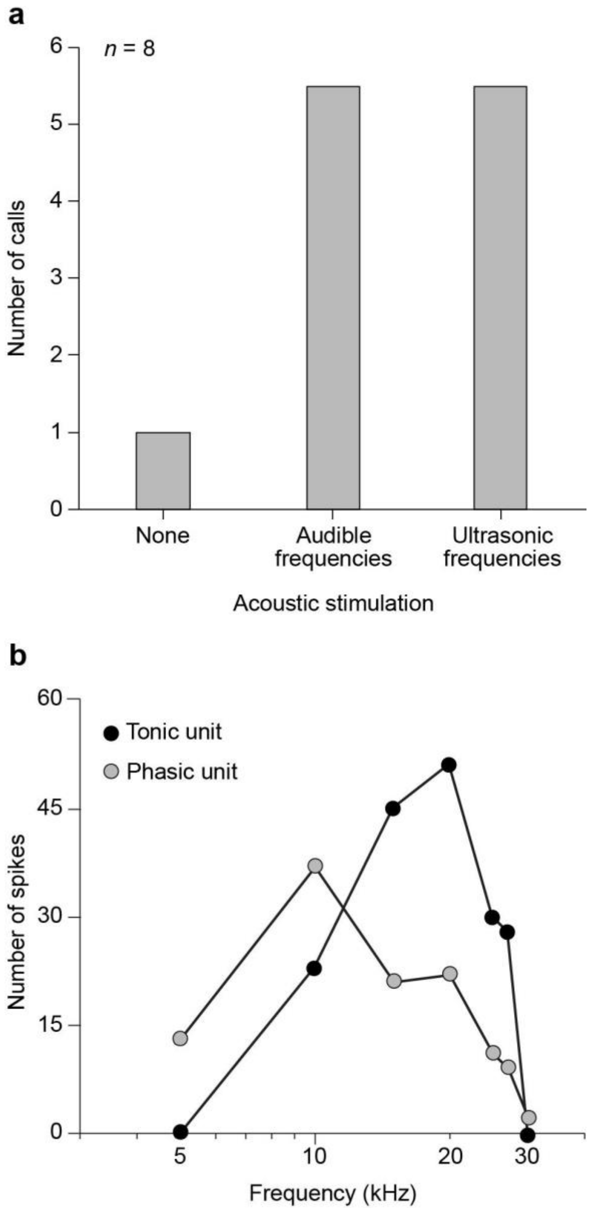

a
n = 8
Number of calls
0
1
2
3
4
5
6
None
Audible frequencies
Ultrasonic frequencies
Acoustic stimulation
b
● Tonic unit
○ Phasic unit
Number of spikes
0
15
30
45
60
5
10
20
30
Frequency (kHz)

Fig. 4

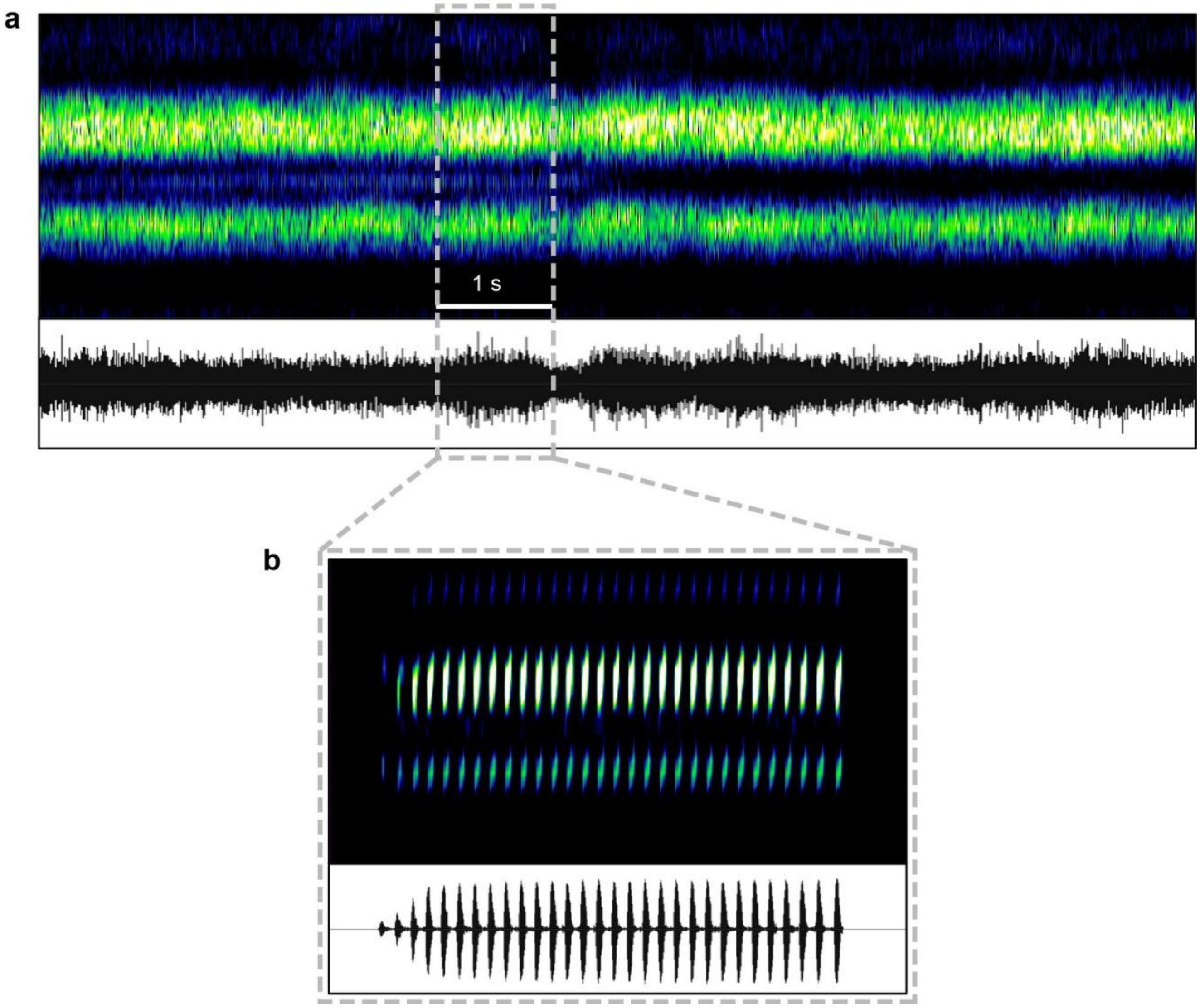


Fig. 5

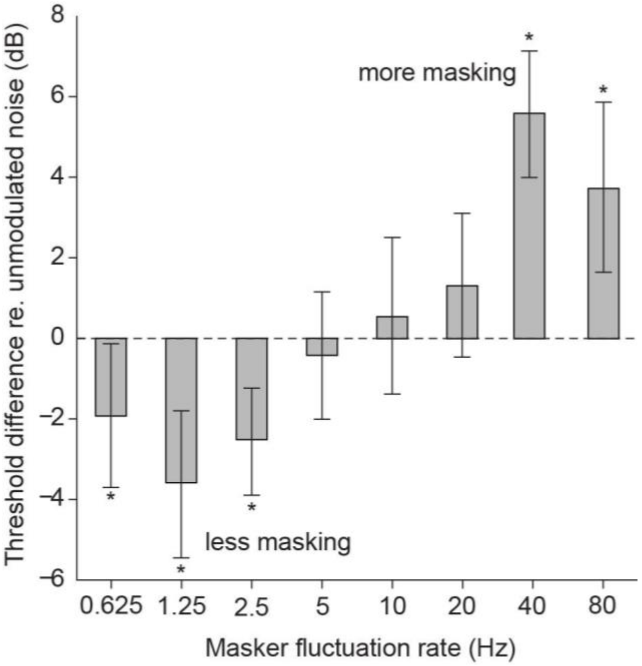

8
6
4
2
0
−2
−4
−6
Threshold difference re. unmodulated noise (dB)
more masking
less masking
*
*
*
*
*
*
0.625
1.25
2.5
5
10
20
40
80
Masker fluctuation rate (Hz)

Fig. 6

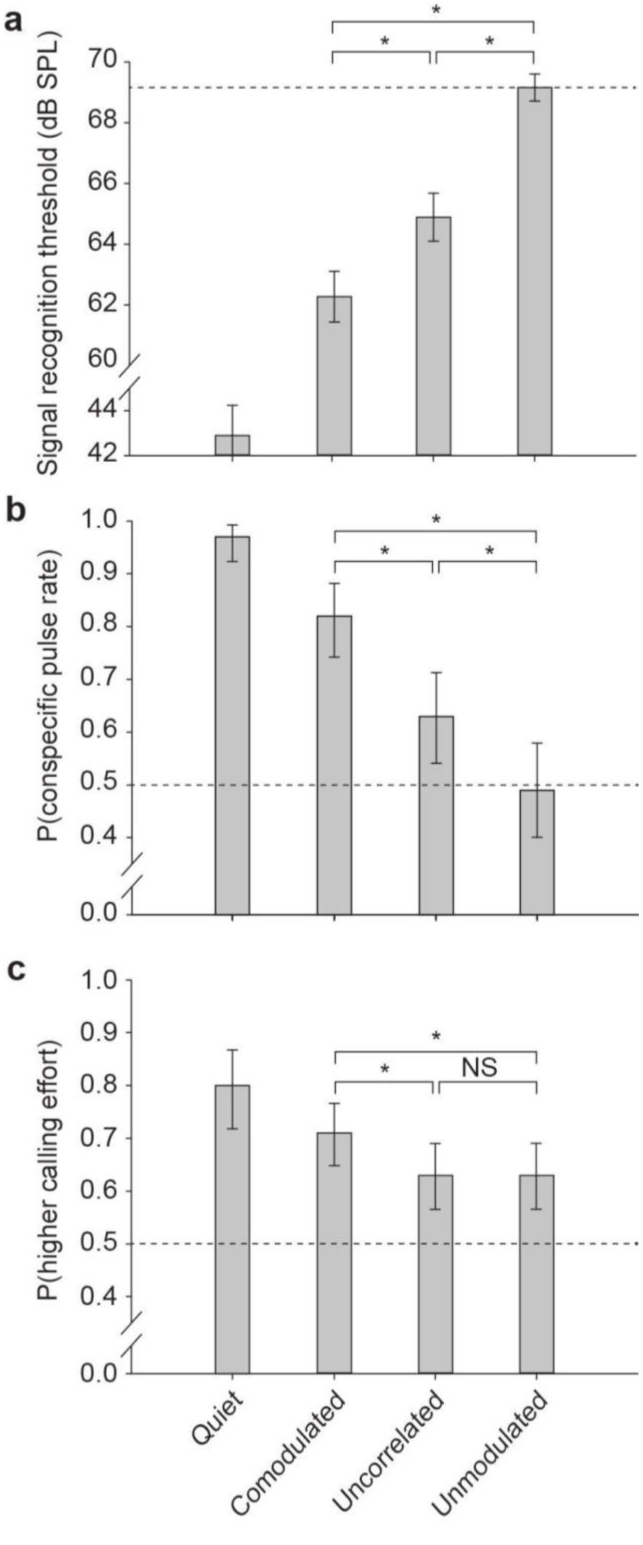

a
Signal recognition threshold (dB SPL)
70
68
66
64
62
60
44
42
*
*
*
b
P(conspecific pulse rate)
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.0
*
*
*
c
P(higher calling effort)
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.0
*
*
NS
Quiet
Comodulated
Uncorrelated
Unmodulated

Fig. 7

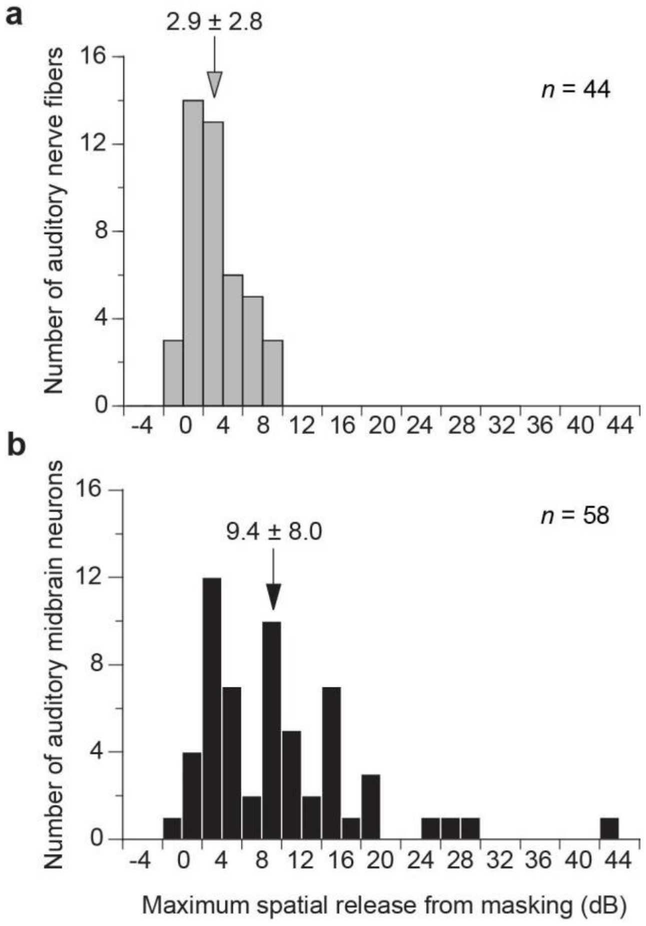

a
2.9 ± 2.8
n = 44
Number of auditory nerve fibers
16
12
8
4
0
-4
0
4
8
12
16
20
24
28
32
36
40
44
b
n = 58
9.4 ± 8.0
Number of auditory midbrain neurons
16
12
8
4
0
-4
0
4
8
12
16
20
24
28
32
36
40
44
Maximum spatial release from masking (dB)

Fig. 8

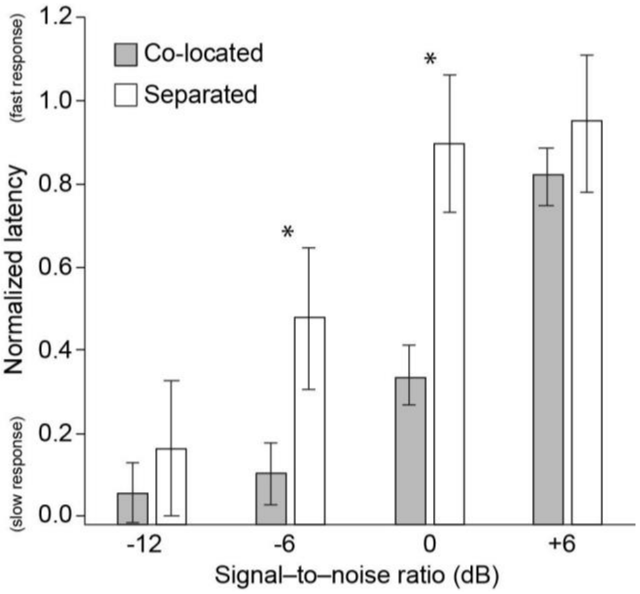
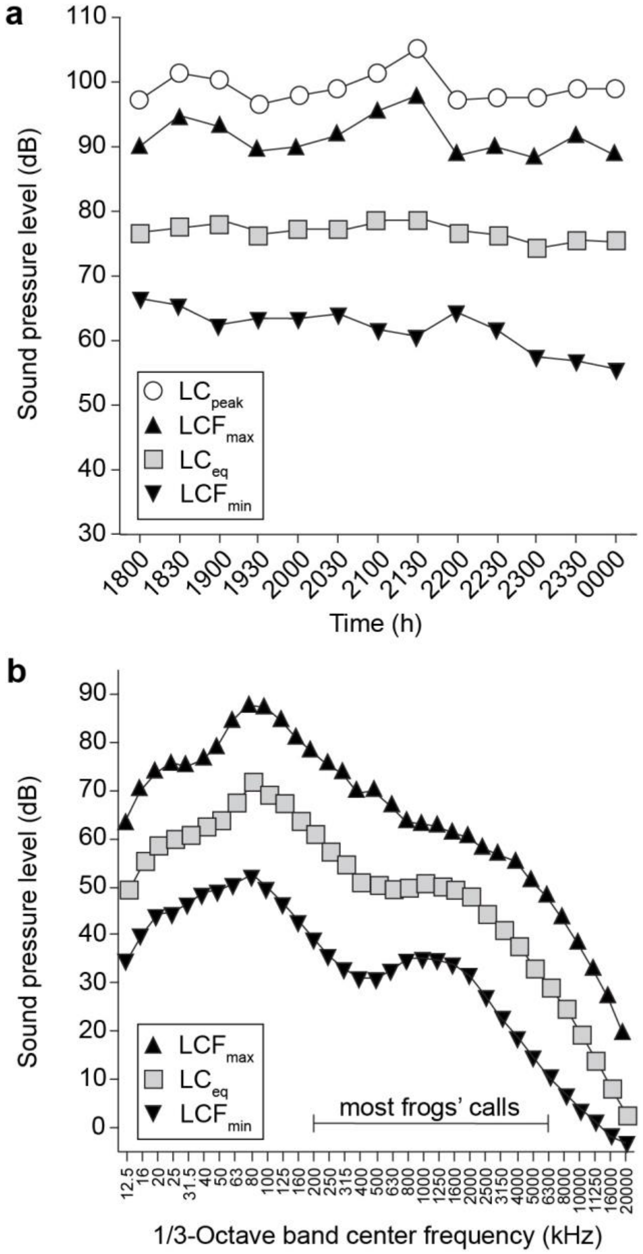

▨ Co-located
☐ Separated
*
*
Normalized latency
(fast response)
(slow response)
1.2
1.0
0.8
0.6
0.4
0.2
0.0
-12
-6
0
+6
Signal–to–noise ratio (dB)

Fig. 9


a
Sound pressure level (dB)
110
100
90
80
70
60
50
40
30
LC$_{peak}$
LCF$_{max}$
LC$_{eq}$
LCF$_{min}$
1800
1830
1900
1930
2000
2030
2100
2130
2200
2230
2300
2330
0000
Time (h)
b
Sound pressure level (dB)
90
80
70
60
50
40
30
20
10
0
LCF$_{max}$
LC$_{eq}$
LCF$_{min}$
most frogs' calls
12.5
16
20
25
31.5
40
50
63
80
100
125
160
200
250
315
400
500
630
800
1000
1250
1600
2000
2500
3150
4000
5000
6300
8000
10000
11250
16000
20000
1/3-Octave band center frequency (kHz)