



Are frog calls relatively difficult to locate by mammalian predators?

Douglas L. Jones¹ · Rama Ratnam²

Received: 25 May 2022 / Revised: 26 October 2022 / Accepted: 8 November 2022 / Published online: 12 December 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

Frogs call in acoustically dense choruses to attract conspecific females. Their calls can potentially reveal their location to predators, many of which are mammals. However, frogs and mammals have very different acoustic receivers and mechanisms for determining sound source direction. We argue that frog calls may have been selected so that they are harder to locate with the direction-finding mechanisms of mammals. We focus on interaural time delay (ITD) estimation using delay-line coincidence detection (place code), and a binaural excitatory/inhibitory (E/I) ITD mechanism found in mammals with small heads (population code). We identify four “strategies” which frogs may employ to exploit the weaknesses of either mechanism. The first two strategies used by the frog confound delay estimation to increase direction ambiguity using highly periodic calls or narrowband calls. The third strategy relies on using short pulses. The E/I mechanism is susceptible to noise with sounds being pulled to the medial plane when signal-to-noise ratio is low. Together, these three strategies compromise both ongoing and onset determination of location using either mechanism. Finally, frogs call in dense choruses using various means for controlling synchrony, maintaining chorus tenure, and abruptly switching off calling, all of which serve to confound location finding. Of these strategies, only chorusing adversely impacts the localization performance of frogs’ acoustic receivers. We illustrate these strategies with an analysis of calls from three different frog species.

Keywords Anuran vocalization · Asymmetric pressure-difference receiver · Sound localization · Mammalian hearing · Interaural time difference

Introduction

Vocally communicating anurans (frogs and toads) congregate in large gatherings around breeding ponds and vocalize in dense choruses to attract females (Gerhardt and Huber 2002). In most frog species, males produce a stereotyped call, and females locate and select a calling male and move towards him using his voice alone, i.e., by phonotaxis (Feng et al. 1976). Given the large number of frogs calling at the same time, the ability to detect the direction from which males are calling (directional hearing) plays a critical role in

anuran reproduction. The middle-ear cavities in anurans are strongly coupled acoustically via the mouth through interaural passages (Feng 1980; Feng and Shofner 1981; Lewis and Narins 1999; Christensen-Dalsgaard 2005, 2011; Ho and Narins 2006; Narins 2016), in a manner that produces about a 10 dB reduction in the pressure gradient across the contralateral eardrum for acoustic sources off to one side (Palmer and Pinder 1984; Jorgensen 1991; Christensen-Dalsgaard 2005, 2011). These receivers therefore act as asymmetric pressure-difference receivers. Each eardrum has a bandpass frequency response and roughly ovoidal directionality. The direction of a source is thus encoded mechanically in the relative interaural level difference between the eardrums, regardless of the detailed structure of the acoustic waveform. Coupled ears are found across the vertebrate taxa and include birds, crocodiles, and lizards. Together with anurans they encode sound direction at the eardrum, although the relative contributions of eardrums and central mechanisms for birds are not fully resolved and may differ across taxa (Christensen-Dalsgaard 2005, 2011).

Handling editor: Peter M. Narins

✉ Rama Ratnam
rama.ratnam@ahduni.edu.in

¹ Electrical and Computer Engineering and the Neuroscience Program, The Grainger College of Engineering, University of Illinois at Urbana-Champaign, IL, Champaign, USA

² Biological and Life Sciences Division and the Psychology Programme, School of Arts and Sciences, Ahmedabad University, Ahmedabad, Gujarat, India

Mammalian ears, on the other hand, constitute two independent pressure receivers which are inherently non-directional. Thus, the mammalian auditory system must rely on central neural comparisons of the signals arriving at the two ears, i.e., binaural comparisons, to estimate sound direction. Mammalian receivers are exceptional in that they are not predominant among vertebrates, where coupled ears are the norm.

Male frogs call at high intensity and frequently so that they can be located by conspecific females. However, these same calls expose their presence and potentially their location to their predators (see Ryan et al. 1981). These predators include snakes, larger frogs, night herons, small nocturnal mammals such as raccoons, weasels, and opossums (a marsupial). Given that mammals exhibit a completely different type of acoustic receiver, it would be advantageous for frogs to evolve calls that are difficult for mammals to locate, while still being locatable by female frogs. This would defeat at least one class of predators. There are few studies on predation risk in chorusing anurans, with the majority being devoted to the túngara frog which is preyed on by the fringe-lipped bat, larger frogs, and at least one terrestrial mammal, the four-eyed opossum (Tuttle and Ryan 1981; Ryan et al. 1981). The risk of predation by the opossum (*Philander opossum*) was lower than the risk from frogs and bats (Ryan et al. 1981) although there is a brief report that suggests that the opossum may be locating the túngara frog by sound alone (Tuttle et al. 1981). Predation of the túngara frog by the fringe-lipped bat is taken up later.

In mammals, the cochlea serves as a frequency-selective filter bank that divides the signal at each ear into many overlapping frequency bands, which are then processed in central auditory pathways. Two largely separate binaural neural processing pathways determine source direction, one based on interaural level differences (ILD) and the other on interaural time differences (ITD). With independent pressure receivers on either side of the head, interaural level differences arise from head shadowing, whereas interaural time differences arise from the finite speed of sound and the longer path length to the contralateral ear. Interaural level differences are negligible for wavelengths exceeding the diameter of the head and are thus significant only at higher frequencies, whereas the interaural time-delay mechanism dominates directional hearing for frequencies below this limit. The calls of the frogs we have studied fall below this frequency limit for the head sizes of small mammals (diameter of a few centimeters) which are likely to prey on frogs, and so this paper focuses on the ITD mechanisms.

In mammals there are two different mechanisms proposed for localizing sounds based on ITD, the dual-delay line mechanism (Jeffress 1948) and the more recent binaural excitation-inhibition mechanism (see Grothe et al. 2010, and reviews below). In the dual-delay line mechanism, the

ITD system localizes the direction of sound arrival, largely independently and in parallel, in each frequency band. At the level of the brainstem, the ITD pathway uses a dual delay-line circuit in the medial superior olive (MSO) that effectually cross-correlates the binaural signals to determine the interaural time-delay with best waveform coincidence. There is evidence for a cross-correlation based coincidence mechanism in the cat (Yin et al. 1987; see Joris et al. 1998 for a brief review) but no clear evidence for a spatial array of cells in the MSO forming a dual-delay line. A spatial array has been convincingly shown in the barn owl, which also use a coincidence mechanism (Carr and Konishi 1990), and chicks (Overholt et al. 1992; Young and Rubel 1983). The dual delay-line system is particularly powerful and effective for localizing low-frequency sounds, for which ILD's are negligible. It nonetheless has drawbacks with respect to pressure-gradient receivers in some situations; namely, it produces ambiguous directional estimates for periodic high frequency sounds (discussed in detail further below). As the ambiguous, i.e., false directions, differ across frequency, the mammalian system can overcome this ambiguity for broadband sounds by integrating the dual-delay line coincidences across frequency; only the true peak appears consistently at the same angle across a broad range of frequencies (Stern et al. 1988). There is some physiological evidence for this mechanism at least in the barn owl, where space-coding neurons in the external nucleus of the midbrain inferior colliculus respond to broad-band sounds tuned to a certain location in space but are less responsive to other features of the sound (Singheiser et al. 2012; Takahashi and Keller 1994; Knudsen and Konishi 1978). A team in Albert Feng's laboratory extended the insights gained from dual-delay line circuits in the barn owl to the stencil-filter concept, which integrates the entire two-dimensional frequency versus delay peak map (the "stencil") against a 2D "stencil filter" that includes the shifted patterns of the false peaks as well. This extension often enhances the detection and localization of multiple simultaneous wide-band sources, such as in cocktail-party environments (Liu et al. 2000).

There is limited evidence for a dual delay line in mammals (for reviews, see Ashida and Carr 2011; Grothe et al. 2010; McAlpine and Grothe 2003) and especially in small mammals which generate small ITDs. Heffner and Heffner (1987) reported that the least weasel (*Mustela nivalis*), a carnivore with an interaural distance similar to that of a mouse, is the smallest carnivore, with a maximum ITD of 76 μ s. However, the localization threshold of the least weasel exceeds that of rodents (Heffner and Heffner, 1987), and is larger than that of mammals with large heads like the horse and cattle (Heffner et al. 2007; Koay et al., 1998). Thus, interaural distance by itself may not solely determine localization acuity. It is likely that carnivores such as cats, dogs, and the least weasel are under greater pressure than

their prey to accurately determine sound source location (see Heffner and Heffner 1987). Nevertheless, a small head size can limit localization acuity and sensitivity as it compresses the azimuthal axis into a small range of naturally occurring ITDs. How then do animals with small heads overcome this limitation?

Recent work suggests that ITD processing in mammals, especially small mammals, may be served by a population of neurons in the MSO. These low-frequency neurons respond to ITDs beyond the range predicted by head-width alone. This is achieved by a balance of excitatory and inhibitory drives from either side to MSO neurons (Harper and McAlpine 2004; McAlpine et al. 2001; Brand et al. 2002; Grothe and Sanes 1994). The inhibitory input (Grothe and Sanes 1994), not considered in dual-delay line coincidence processing, is crucial to ITD tuning and is glycinergic (Brand et al. 2002; see Brughera et al. 1996 for a computational model). Inhibition shifts the peak of the ITD function outside the normal range of ITDs, typically at an interaural phase difference of 45° , so that the most sensitive, monotonic, portion of the ITD function is placed within the physiological range. The balance of activity on either side can then provide an estimate of the ITD and hence azimuth. Thus, the excitation/inhibition or E/I binaural ITD model can solve the conundrum of how mammals with small heads and hence small ITDs localize sound at low frequencies.

Ongoing ITD disparities serve as important cues for localizing sounds at low frequencies especially under closed-loop conditions where the organism can integrate over long time scales. However, sound localization is possible under open-loop conditions by timing onset disparities. This is useful in reverberant environments where the direct sound reaches the ears first (the “first wavefront”) and is followed by successively weaker copies of the sound originating from spurious locations. Reflections from different directions can confound the dual delay-line estimate of the true sound-source direction. The mammalian auditory system suppresses reflections arriving about a millisecond or so after the first wavefront, presumably by using an onset gating mechanism. Direction of the leading sound (the true source) is determined from the first few milliseconds of sound following a sudden increase in amplitude. A process called the “precedence effect” then suppresses the localization mechanism to lagging sounds, in humans, for about 5 milliseconds (Litovsky et al. 1999). The precedence effect works similarly in small mammals such as the cat (Tollin and Yin 2003) and the ferret (Tolnai et al. 2014).

More generally, onset detection may be critical in several signal detection tasks not restricted to listening in reverberant environments. The cocktail-party effect (Cherry 1953) is among the most important listening tasks in real environments. It has been widely studied as spatially mediated release from masking in humans (see Saberi et al. 1991;

Plomp and Mimpen 1981) and in anurans (see Bee 2007, 2008; Schwartz and Gerhardt 1989). At the neural level, several studies have shown that onset detection by means of phasic neurons may be a basic mechanism underlying spatially mediated release from masking (Feng and Schul 2007; Lin and Feng 2001; Feng and Ratnam 2000; Ratnam and Feng 1998). These neurons are effective when the angle of attack is steep, i.e., the onset is abrupt, providing a cue for rapid estimation of the direction of a sound source. Thus, onset detection can facilitate spatially mediated release from masking and form the basis for listening in cocktail party environments and frog choruses. Mammalian predators that prey on frogs calling in a chorus can potentially use both onset and ongoing interaural disparities to locate frogs. We discuss ways in which frogs structure their calls to make it harder for mammals to detect onsets or make effective use of ongoing disparities to improve their localization performance.

In the following study, we take a detailed look at the acoustic structure of frog vocal signals and propose four “strategies” that make it difficult for mammals to localize frog calls. The first two strategies defeat the normal mechanisms of ITD processing and directional hearing in mammals so as to make spatial location ambiguous. The third strategy proposes that pulse structure and timing in the calls of frogs make it harder to detect signal onset and reject interference and noise. Finally, we touch upon a fourth strategy, namely the benefits of calling in a chorus.

Methods

Vocalizations from three species of frogs are reported in this work: (1) The cricket tree frog (*Acris crepitans*), (2) the túngara frog, *Engystomops pustulosus* (formerly *Physalaemus pustulosus*), and (3) the gray treefrog, *Dryophytes versicolor* (formerly *Hyla versicolor*). The cricket frog calls were recorded at the Cibolo Nature Center in Boerne, Texas, on the evening of March 21, 2007 and were available at a sampling rate of 20 kHz and 24-bit resolution. These calls were analyzed for chorus structure in an earlier study (Jones and Ratnam 2009). The túngara frog data were downloaded from a publicly accessible website (Hermans 2019) and was available as a 16-bit WAV file at a sampling rate of 44.1 kHz. Other details of this recording are not known. The gray tree frog recordings were obtained from Dr. Joseph A. Dellinger (Houston, TX) who made the recordings near a roadside ditch at Brazos Bend State Park (Fort Bend County, TX) on March 15, 2014, using the internal cardioid microphone of a Zoom H4N digital audio recorder. These calls were available as 16-bit WAV files sampled at 48 kHz. In all cases, only a single channel from each recording was considered, even if

multi-microphone data were available. Other channels were discarded and do not appear in the analysis.

Each call was bandpass filtered with a Sinc (rectangular) filter to eliminate noise and other signals unrelated to the call. The high-pass and low-pass frequencies for the filter were: (1) 2.7 and 4 kHz (cricket frog), (2) 0.4 and 4.4 kHz (túngara frog), and (3) 1.55 and 2.8 kHz (gray tree frog). The frequency spectrum and autocorrelation were estimated from the filtered call using the discrete Fourier transform (DFT). A single filtered call from each of the three species was considered for analysis.

To simulate a mammalian cochlear filter bank, we used a bank of constant-Q bandpass IIR filters (a cascade of two second-order Butterworth filters) which are equally spaced in log frequency with a density of 12 per octave, and quality factors (Q) of about 9. This filter bank approximates those found in the mammalian auditory system. For example, Glasberg and Moore (1990) report Q's ranging from 6 at very low frequency up to 9.6 at higher frequencies. Oxenham and Shera (2003) argue for Q ranging from 9 to 12 in humans, but also noted that Q values for humans are higher than those of most other mammals, such as cats. We selected a Q of 9 for this study.

“Binaural” signals with only ITD (but no ILD) cues were generated by delaying the outputs from the filter bank. There was no criterion used in selecting the delay except to illustrate the binaural ITD-based localization schemes used here, and so the delay resulted in an arbitrary angle of arrival. Independent bandpass Gaussian noise was added to each binaural channel with the noise scaled relative to the peak of the call. For quiet conditions, the noise power was adjusted so that the signal-to-noise ratio (SNR) was 60 dB or greater. Noise in each binaural channel was generally fixed at 20 dB SNR, but in some cases we assessed the efficacy of both dual delay line coincidence and a small mammal binaural E/I model at an additional 6 dB and 0 dB SNR. Head diameter was assumed to be 0.15 m for the dual delay line model and 0.01 m for the binaural E/I model, and the velocity of sound 342 m/s.

We implemented two models of binaural processing: (1) the Jeffress dual delay line model, and (2) the binaural E/I model. In the first model, the noisy binaural signal was input to a dual delay line, which was implemented as follows: At each center frequency, and for each corresponding bandpass filter pair, the left and right channels are correlated (Fig. 1c) with delays corresponding to 1° increments from 0° to 180° ($0-\pi$). The fractional sample delays were implemented in the Fourier domain by a linear phase shift of the DFT. This produces a two-dimensional coincidence map of frequency as a function of azimuth angle (Fig. 1d). Local maxima of the coincidence map in each frequency band are set to 1, and all other azimuth angles for that frequency are set to zero, thereby producing a “stencil” (Fig. 1e). This stencil

is summed (filtered) across frequency for each angle, and averaged with the left and right neighboring angles, thereby producing a summed stencil filter output (Stern et al. 1988) (Fig. 1f). For simplicity, we omit the full two-dimensional stencil filter including the summation along the ambiguity arcs seen in Fig. 1d and e, which provide benefit mostly for localizing multiple simultaneous broadband sources (Liu et al. 2000). The largest peak of the summed stencil is the azimuthal location estimate.

The second model implements a signal processing algorithm reflecting the current understanding of small mammal binaural ITD processing (McAlpine and Grothe 2003; Grothe et al. 2010; Harper and McAlpine 2004; Brand et al. 2002; McAlpine et al. 2001; See also a similar model by Brughera et al. 1996). On one side, let us say the left, the inhibitory and excitatory ipsilateral and contralateral inputs to neurons in the medial superior olive (MSO) are relatively delayed so as to create a “sandwich” of inhibitory “bread” surrounding an excitatory center as shown in Fig. 2 for a 1 cm head-width, for three different ITD conditions: contralateral lagging (negative ITD, Fig. 2a), broadside (medial plane, zero ITD, Fig. 2b), and contralateral leading (positive ITD, Fig. 2c). All panels in Fig. 2 show the responses of the four inputs consisting of three periods of a half-wave rectified sinusoid. While the relative strengths and timings of inhibition and excitation are not known (Grothe et al. 2010), we assume (see Brand et al. 2002) that ipsilateral inhibition (green) lags ipsilateral excitation (blue) and contralateral excitation (orange), and that contralateral inhibition (red) leads all the other responses. This creates a sandwich with the inhibitory inputs acting as slices of bread on the outside, and the excitatory inputs acting as the filling. At zero ITD (broadside, Fig. 2b), the delays are arranged such that the excitatory inputs are about half uncovered from the surrounding inhibitions. The total output is computed as the ongoing half-wave rectified sum of the two excitatory inputs minus the sum of the two inhibitory inputs (thick black lines), the mean of which is computed over some time interval to produce the rate equivalent response (Fig. 3a, filled blue circle marked ‘b’). A similar total is computed for the output on the right side (Fig. 3a, red circle overlaid by blue at ‘b’). With positive ITD (contralateral leading), the increasing separation between the surrounding inhibitions (see Fig. 2c) increases the excitatory response up to a certain point outside the physically realizable ITD before falling again as inhibitions from subsequent periods come into play. The instantaneous half-wave rectified output on the left side (Fig. 2c, thick black lines) results in increased mean output (Fig. 3a, filled blue circle marked ‘c’). At negative ITDs (Fig. 2a) the sandwich closes, and the larger inhibitions overlap and suppress the excitatory response, producing a diminished instantaneous half-wave

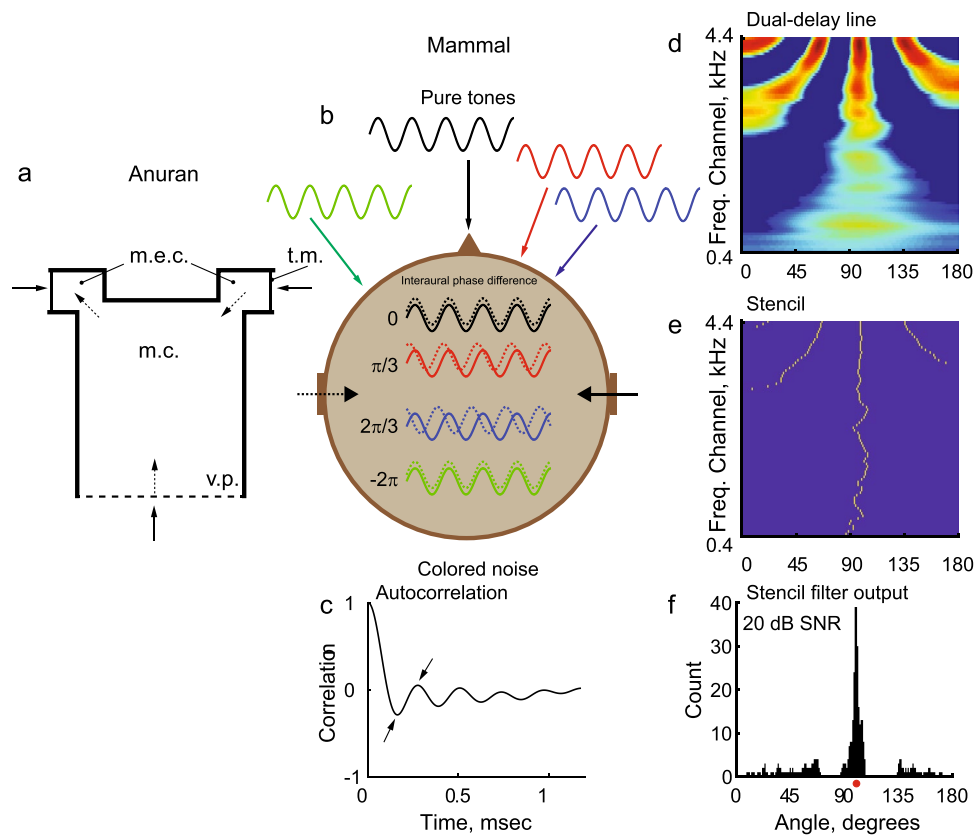


Fig. 1 Receiver characteristics of frogs and direction estimation in mammals. **a** The asymmetrical pressure-difference receiver of the frog illustrating the multiple acoustic pathways to each ear. The right tympanic membrane (t. m.) receives direct sound on the external surface as does the left. The input from the right t. m. is transmitted through the middle ear cavity (m. e. c.), the mouth cavity (m. c.) to the internal surface of the left t. m. Another indirect pathway originates in the vocal pouch (v. c.) and goes through the m. c. to the internal surface of the left t. m. Similarly for the pathways to the right ear. (Adapted and redrawn from Feng and Shofner 1981). **b** The mammalian ears are independent pressure receivers. Shown are four pure tones of the same frequency originating from different azimuths. The source at the midline (black) arrives in phase at the two ears (left: dotted, right: solid). The sources to the right arrive at different phases with right leading (red, phase = $\pi/3$; blue, phase = $2\pi/3$). The source

to the left (green) is left leading but arrives at the same phase (green, phase = -2π) and is impossible to disambiguate from the source at the midline. **c** Correlation function of a white noise source placed just a little to the right of the midline. Noise was bandpass filtered so that its frequencies lie between 0.4 and 4.4 kHz. The secondary peaks (arrows) are greatly attenuated allowing the primary peak to be potentially unambiguously detected by a dual delay-line cross-correlator shown in **d**. However, the dual delay line shows ambiguities in source location (abscissa) at higher frequencies (ordinate). True source location is just to the right of 90°. **e** A stencil finds the peak (coincidence) in each frequency band and scores it as the source location at that frequency. **f** At each azimuth the number of peaks across frequency are summed to produce a histogram of peak counts at that azimuth (output of the stencil filter). False coincidences sum incoherently whereas the true location provides the strongest peak.

rectified response (Fig. 2a, thick black line, marked by arrow) and diminished mean total output (Fig. 3a, filled blue circle marked 'a'). The responses shown in Fig. 3a are similar to responses recorded from neurons of the MSO of small mammals (Brand et al. 2002). A normalized difference metric, defined as the difference of the left and right responses divided by their sum, produces a nearly linear metric over the feasible ITD range, from which the ITD (and the equivalent forward azimuth) can be determined (Fig. 3b, the filled circles and letters correspond to the total mean output of Fig. 2a and b, and c, respectively). These simulation results are consistent with recent

research and models of ITD processing in small mammals (Harper et al. 2014; Grothe et al. 2010; Harper and McAlpine 2004; McAlpine and Grothe 2003; Brand et al., 2002; McAlpine et al. 2001).

Our model for a 1 cm head (Figs. 2 and 3) delays the leading contralateral inhibition by one cycle period ($1/\text{center frequency}$) \pm the ITD, with the contralateral excitation delayed by an additional 0.35 of the cycle period. The ipsilateral excitation is delayed by 1.4 periods, with an additional delay of 0.2 periods to the ipsilateral inhibition. The relative delays thus scale with the center frequency as do those reported in the literature (McAlpine et al. 2001), producing a similar

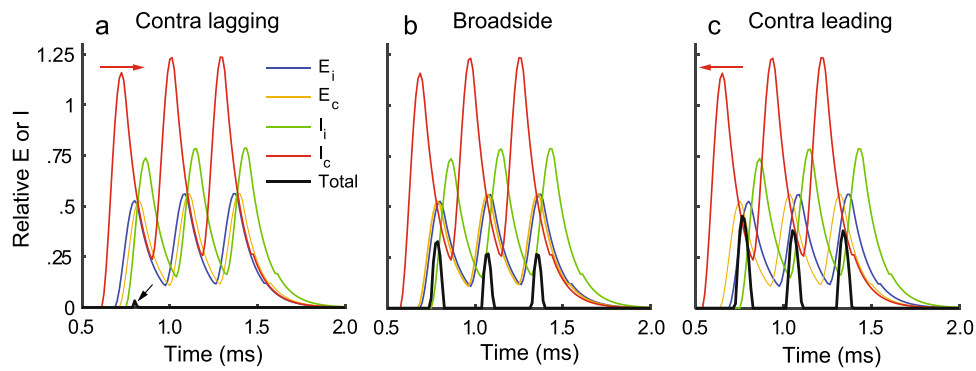


Fig. 2 The “Inhibitory sandwich” implementation of the excitatory/inhibitory (E/I) model of ITD processing in the medial superior olive (MSO) neuron of a hypothetical small mammal (head-width = 1 cm). Each panel depicts model response for excitatory (E) and inhibitory (I) drives from contralateral (c) and ipsilateral (i) sides with contralateral inhibition (I_c red) **a** lagging, **b** broadside ($ITD=0$), and **c** leading. Relative magnitude (ordinate) is arbitrary scale. The onset and time course of ipsilateral drives, E_i (blue) and I_i (green), are fixed in time

with contralateral drive (E_c and I_c) timing adjusted according to ITD (red arrows show shift in the curves relative to broadside). The timings are created so that the excitatory drives are sandwiched between the inhibitory drives. The half-wave rectified output of the total drive (excitation – inhibition) is the net output of the MSO neuron (black). Total response diminishes with increasing contralateral lag (black arrow, **a**)

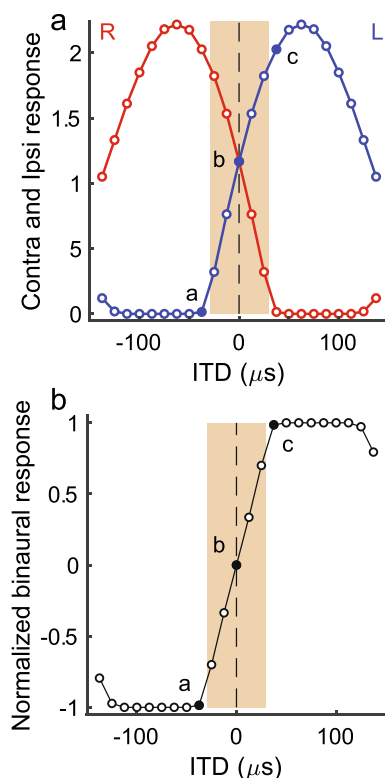


Fig. 3 Output of model MSO neuron shown in Fig. 2 across all ITDs. For each panel in Fig. 2, the area under the total response curve (black, Fig. 2) is shown (ordinate) as a function of ITD (abscissa). **a** Response of hypothetical neurons on the left (blue) and right (red) side. Positive ITD is contralateral leading. The filled circles (**a**, **b**, and **c**) correspond to Fig. 2a and b, and c. **b** Total response curves (from panel a) are summed and normalized by peak value

response pattern at all frequencies. These responses were filtered with first-order low-pass synaptic response filters with inhibitory (glycinergic) and excitatory (glutamatergic) time constants of 0.1 ms (Brand et al. 2002). The relative strengths of the excitations and inhibitions are 1.0 (both ipsilateral and contralateral excitations), -2.0 (contralateral inhibition), and -1.4 (ipsilateral inhibition) at a center frequency of 3.5 kHz, and -3.8 (contralateral inhibition) at a center frequency of 2 kHz.

It is important to note that these algorithms represent the ideal overall responses of very large populations of binaural neurons at the same frequency with systematic variations in the relative delays of the model components and should be considered as upper bounds on the performance of a finite population of spike-based model neurons.

Strategy I: time-delay ambiguity via periodically repeating waveforms

Dual-delay-line cross-correlation locates sources by finding the delay corresponding to maximum coincidence between the sound received at the left and right ears. The delay maps systematically to a corresponding direction of arrival based on the effective acoustic path length between the two ears and the speed of sound, c . Under certain conditions, and for a source in front of a head of diameter $2L$, angle of arrival θ , the interaural time difference (ITD) is given by Woodworth’s model as (Woodworth 1938; see also Aaronson and Hartmann 2014):

$$ITD = \frac{L}{c}(\theta - \sin\theta)$$

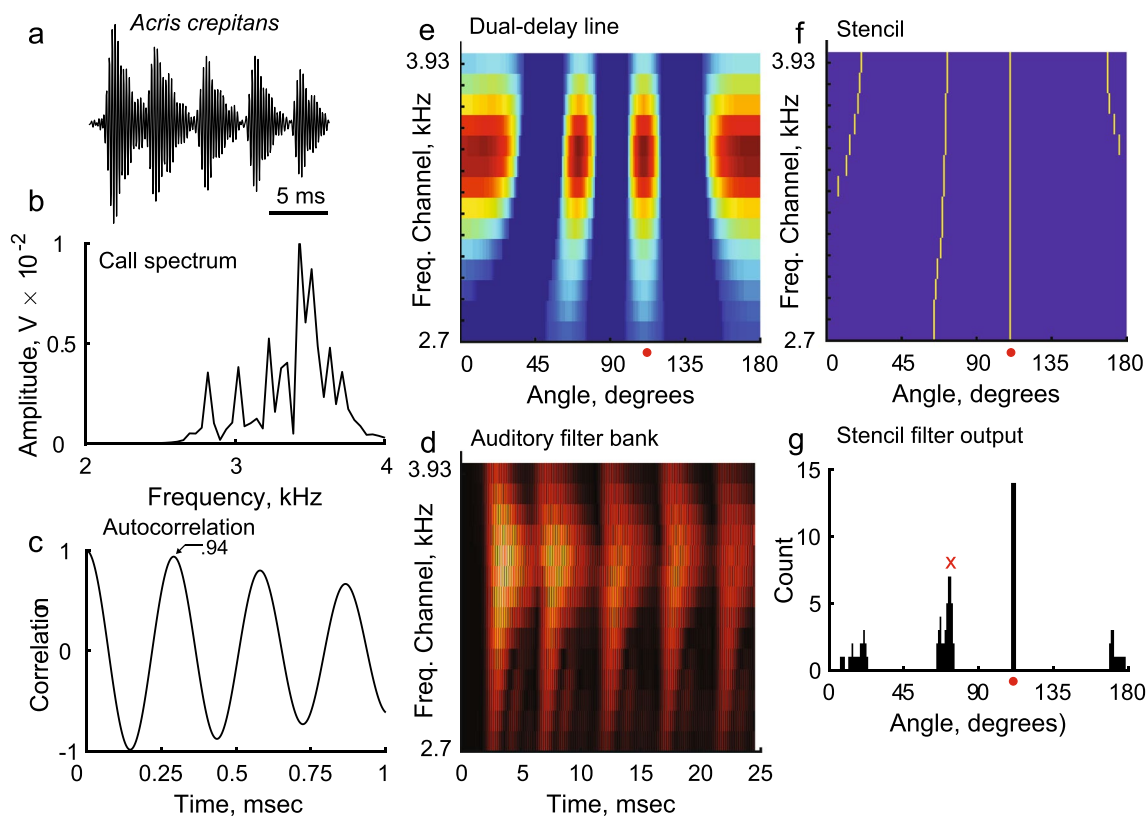


Fig. 4 Localization of a single call of the cricket frog (*Acris crepitans*) using mammalian ITD (interaural time difference) processing. **a** A single call of the cricket frog, **b** the frequency spectrum of the call, and **c** the autocorrelation function showing the nearly periodic nature of the call, and the prominent second peak ($r=0.94$). **d** The output of a model mammalian auditory constant-Q filterbank with 12 channels/octave (abscissa: time; ordinate: frequency). Input to filter

is the call, in quiet conditions (60 dB SNR). The five pulses of the call are visible (see **a**). **e** The output of the dual delay line showing the smearing of the coincidence map (filled red circle: true location). **f** Stencil showing the true and ambiguous directions. Note the prominent false peak to the left of the true location. **g** Stencil filter output shows the true location, and a prominent ambiguous location (x, red) even though the SNR is very high

Figure 1b illustrates the effect of the interaural time delay for simple sinusoidal signals arriving from different directions and the resulting time offsets as input to the dual delay lines. Note that for periodic signals such as infinite-duration sinusoids, delays corresponding to integer multiples of a period result in perfect phase alignment, with dual delay line input waveforms identical to a source from directly in front of the animal. This creates an ambiguity in the direction of arrival that cannot be resolved by any time-delay estimation method, including the dual delay line-based coincidence detector model. We note that the asymmetrical pressure-gradient receiver of the frogs' ears (Fig. 1a) works on different physical principles and does not suffer this ambiguity for periodic inputs.

According to Woodworth's formula, the maximum interaural time-delay occurs for a source directly to the side of the head:

$$\text{ITD} = \frac{\pi L}{c}$$

The corresponding maximum wavelength (λ_{\max}) and frequency (f_{\max}) are $\lambda_{\min} = \text{ITD}_{\max} c = \pi L$ and $f_{\max} = 1/\text{ITD}_{\max}$, respectively. Below this frequency there is in principle no periodicity ambiguity, although the secondary acoustic path, which is the longer way around the head, extends the range of ambiguity significantly (Aaronson and Hartmann 2014). The mammalian brain encodes ITDs several times larger than this lower limit (Grothe et al. 2010), perhaps enhancing the performance of a stencil filter.

Waveforms with multiple large cross-correlation peaks at delays within the physically plausible range for that animal can confound the directional estimate, because the animal cannot determine which peak corresponds to the true source direction. In this respect, pure tones (sinusoids) are particularly difficult to localize, because the cross-correlation exactly repeats once every cycle (2π phase shift). For example, for an effective head radius of 3 cm, a pure tone at a frequency of 3640 Hz directly to the right will have an identical cross-correlation peak for a source to the left. That

is, the animal cannot determine whether a periodic source is located to its right side or to the left.

The difficulty in localizing frog calls using cross-correlation, i.e., dual-delay line estimation, is illustrated in Fig. 4 with the call of a cricket frog (*Acris crepitans*). Depicted are the call of the frog (Fig. 4a), its amplitude spectrum (Fig. 4b), and the normalized correlation function of this call (Fig. 4c). The correlation corresponds to the output of a wideband, i.e., single channel, dual delay line. The true location is at time zero. The time delay to the closest, and largest, secondary peak (a false location) in the cross-correlation is 0.29 ms, with a correlation value which is 94% (−0.27 dB) of the true peak. This will be discussed further below.

To determine localization performance, a synthetic binaural signal was created by delaying the mono-channel recording (Fig. 4a) by three samples in the left ear channel relative to the right ear channel. This produced a virtual sound source at about 110° (i.e., about 20° from the midline, right ear leading). Sound to each ear was filtered using a constant-Q filter bank (see Methods, and output shown in Fig. 4d). The left and right filter-bank outputs were run through Jeffress's model of a dual delay-line for the mammalian auditory system (Jeffress 1948). Figure 4e is a coincidence map which shows the topographic dual delay-line output for each frequency band (ordinate) as a function of the direction-of-arrival (abscissa). A “stencil” which extracts the local maxima in each frequency band was applied to the coincidence map (Fig. 4f). These local maxima provide estimates of the sound location from a given frequency band. The stencils are summed across frequency bands (along the y-axis) and smoothed using an azimuthal window of 3 degrees to produce the summed coincidence detection map (Fig. 4g). The summed coincidence detection map indicates the directions of acoustic sources in the azimuthal plane (Liu et al. 2000). Additional “false” peaks are noticeable at separations of roughly 45°–50° from the true azimuth of about 110°.

Realistic acoustic environments including frog choruses are generally noisy, with interfering sounds originating from conspecific and heterospecific frogs, and other biotic and abiotic sources. Figure 4e and f, and g were determined under high signal-to-noise ratios (SNRs) of 60 dB SNR (see Methods), i.e., almost quiet conditions. The performance of the coincidence detectors, stencil filters, and summed coincidence detectors under noisy conditions for the cricket frog call are shown for 20 dB SNR (Fig. 5a–c), 6 dB SNR (Fig. 5d–f), and 0 dB SNR (Fig. 5g–i). The source direction ambiguity increases as SNR decreases.

The presence of spurious peaks in Figs. 4 g and 5 c and i are a consequence of the strong periodicity of the call (not necessarily narrow band). Figure 4c shows the wideband correlation function. As noted above, the time delay to the closest, and largest, secondary peak is 0.29 ms, with a correlation value which is 94% (−0.27 dB) of the true peak.

To put this in perspective, at a relatively high SNR of 12 dB, the average power level of the noisy false cross-correlation peak would equal that of the true peak of the clean signal, rendering a false peak larger than the true peak a very frequent occurrence. This makes it impossible for a dual delay line to reliably distinguish the true source direction from that corresponding to a secondary correlation peak for effective interaural distances of 6 cm or more. In addition to larger mammals, this might also suggest why barn owls, which have excellent dual-delay-line based directional hearing, are not known to typically take frogs.

The small-mammal binaural hearing model does not suffer from phase-wrap ambiguities for small heads at these frequencies. However, the excitation/inhibition structure is inherently differential and depends on accurately estimating a small quantity by subtracting larger quantities and is thus inherently sensitive to noise. Figure 6 shows the calculated variability in response of the hypothetical MSO neuron (depicted in Figs. 2 and 3) at SNRs of 20 dB (Fig. 6a), 6 dB (Fig. 6b), and 0 dB (Fig. 6c). The normalized difference metric shown in Fig. 3b is calculated for the 4th pulse of the cricket frog call at five different angles (−59°, −23°, 0°, 23°, and 59°) within a cochlear band centered at 3500 Hz, with 5th, 50th, and 95th percentile error bars (red). At 20 dB SNR (Fig. 6a), the variance of the estimates is significant, and the outer angles are beginning to show a significant bias towards broadside. At 6 dB SNR (Fig. 6b), the estimates are severely biased and have much greater variance, and by 0 dB SNR (Fig. 6c), all angles return the same distribution, and essentially no information about the source direction can be recovered. Subsequent sections examine several “strategies” by which frog calls minimize the SNR available to the small mammal binaural mechanism for directional estimation.

Strategy II: narrowband calls

Periodic signals induce periodicities in the coincidence metrics used by time-delay detectors, thereby creating unresolvable directional ambiguities (Fig. 1d and e). However, the angles of arrival which correspond to “false” peaks will vary with the frequency (Fig. 1e). Directional ambiguities for broadband signals can be resolved by integrating these coincidences across frequency, via a stencil filter or some similar mechanism (Fig. 1f). The true angle of arrival coincides in every band, whereas the ambiguous false coincidences smear across angle when integrated across frequency, thereby leaving a single large peak at the correct angle in the stencil filter output (Liu et al. 2000). There is some support for these ideas, although the strongest evidence comes from a non-mammalian species. In the barn owl, phase ambiguity appears to be resolved in the external nucleus of the inferior colliculus (Fujita and

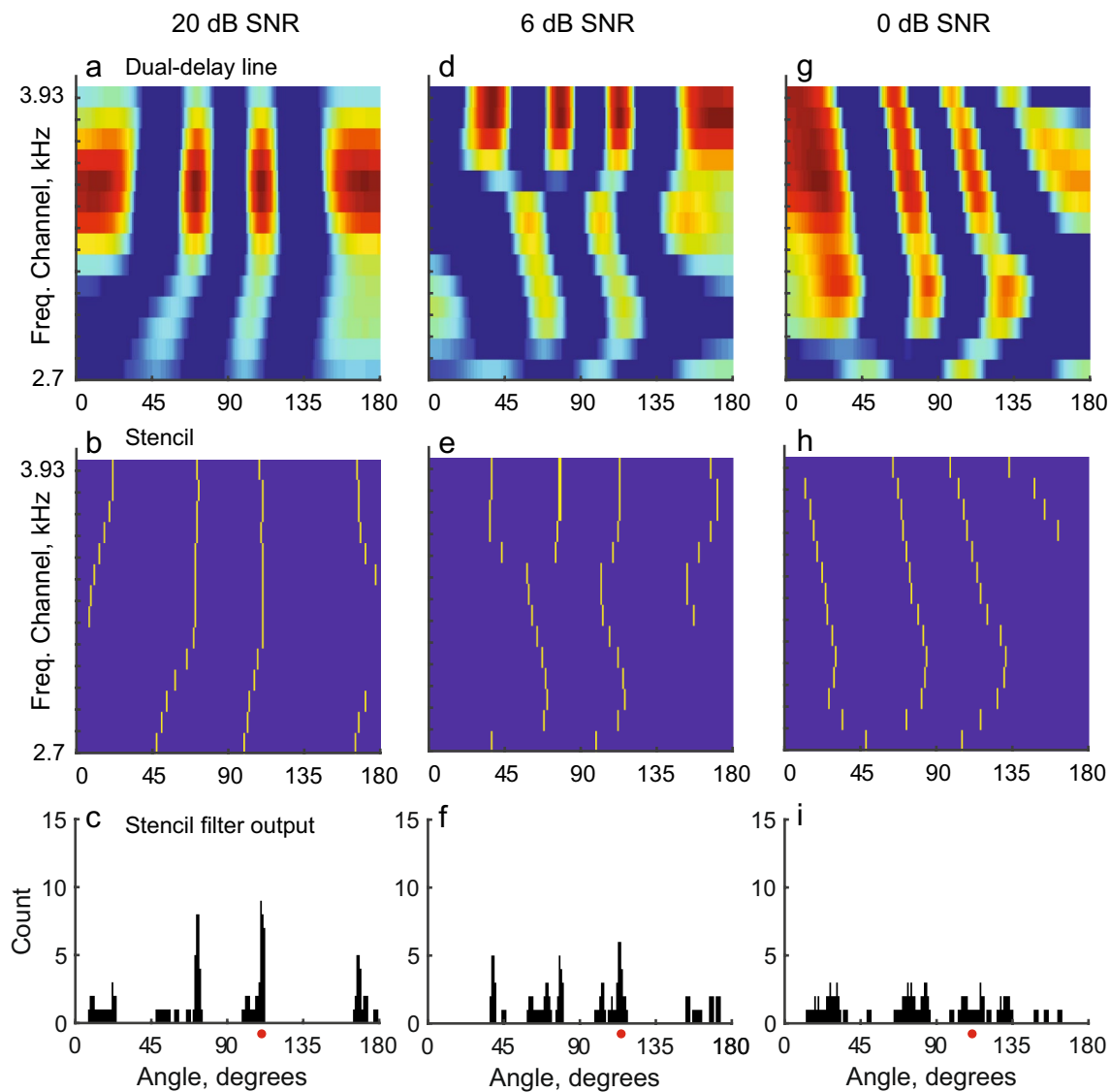


Fig. 5 Dual delay-line processing of cricket frog call shown in Fig. 2 for three different SNRs: 20 dB SNR (a–c), 6 dB SNR (d–f), and 0 dB SNR (g–i). The panels at each SNR follow the description for Fig. 2e–g. With reduced SNR, the location estimate become

degraded. It should be noted that frog choruses are dense and loud and operate under low SNR conditions. Remaining figures are at 20 dB SNR

Konishi 1991) where neurons respond to interaural time disparities independently of stimulus frequency (Knudsen and Konishi 1978). These neurons may be integrating information across the tonotopic axis, in the same way as summation using a stencil filter. It is possible that central mechanisms at the level of the mammalian midbrain may contribute to the resolution of phase ambiguity in a similar way.

This powerful mechanism is most easily defeated by producing calls confined to a narrow band of frequencies. An infinite-duration, pure sinusoid is the only signal with zero bandwidth. It has been mathematically proven that the product of the time duration and the bandwidth of any signal

equals or exceeds a positive, finite value (Gabor 1946). Frog calls, which are of finite duration, must thus extend across some range of frequency. However, for the purpose of confounding mammalian predators, the bandwidth need be no narrower than that of the cochlear filters, which have a quality factor (center frequency to bandwidth ratio) or Q of about 9–13.

Radio engineers have known for over a century that bandwidth-efficient signals can be constructed by amplitude-modulating a sinusoidal “carrier” with a low-pass “envelope” of minimal frequency extent. Discontinuities, rapid changes in amplitude, and sign changes are all known to greatly expand the bandwidth and thus, reduce ambiguity

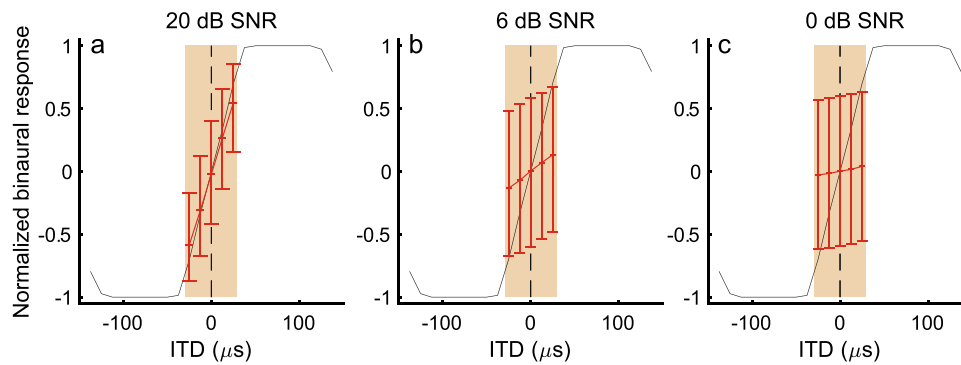


Fig. 6 Localization errors resulting from bias (red line) and variability (red whiskers) due to varying SNR. **a** 20 dB SNR, **b** 6 dB SNR, and **c** 0 dB SNR. The noisy estimates are plotted against the noise-free (high SNR) estimate (black line). Curve follows Fig. 2b. The

whiskers show 5th, 50th, and 95th percentiles of ITD estimates. Note the increase in estimation bias at the lateral angles, which bring the source to the center. At low SNRs the estimates become unusable

in estimating true location. We would therefore expect stealthy frog calls to exhibit smooth, slowly rising and falling envelopes modulating a sinusoidal carrier. Each pulse in the cricket frog call (Fig. 4a) displays these low-bandwidth characteristics, and the spectrum (Fig. 4b) reveals a narrow spectral peak with a quality factor exceeding 10, or finer than the resolution of the mammalian ear. Even at high SNRs, the stencil map (Fig. 4f) shows that the coincidence corresponding to the false location (Fig. 4g, red x) hardly disperses across angle (Fig. 4e and f), and fails to resolve the directional ambiguity of the cricket frog's call. In real choruses where SNR is even lower, even modest noise or chorus interference will lead to increased ambiguity (Fig. 5). The authors can testify to the difficulty in locating vocalizing cricket frogs. Just by ear alone, there appear to be more frogs than there actually are. Only multiple microphones with a large aperture could correctly locate and extract each caller (Jones et al. 2014; Jones and Ratnam 2009).

We further illustrate this strategy with an analysis of the call of a gray treefrog, *Dryophytes versicolor*. A single call of the gray tree frog (Fig. 7a) consists of a sequence of pulses. The depicted call has 11 pulses, the second of which (marked b) is expanded and shown in Fig. 7b. We analyzed this second pulse with a dual-delay line for determining localization performance (Fig. 8). Except where noted, descriptions and methods follow those for the cricket frog and Fig. 4. Depicted are a single pulse from the call (Fig. 8a), and the spectrum of the pulse (Fig. 8b) estimated after bandpass filtering the entire call between 1.55 and 2.8 kHz. The Q factor (at half power) is about 30, indicating a sharply tuned spectrum. The correlation function of the pulse (Fig. 8c) shows that the time delay to the closest, and largest, secondary peak in the cross-correlation is 0.462 ms, with a correlation value which is 96% (−0.18 dB) of the true peak. A second channel was created by delaying the call by three samples (48 kHz sample rate). The two channels

were run through identical constant Q filter banks (Fig. 8d). Noise was added to produce a SNR of 20 dB and the noisy filter-bank outputs were passed through a dual delay-line to produce a coincidence map (Fig. 8e), local maxima were extracted from the coincidence map to produce a stencil (Fig. 8f), and the stencil was summed across frequency to produce the summed coincidence map indicating the directions of acoustic sources in the environment (Fig. 8g). The

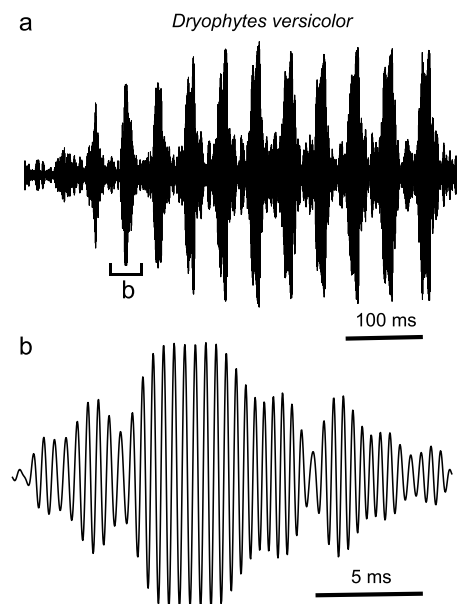


Fig. 7 The call of a gray treefrog, *Dryophytes versicolor* (formerly *Hyla versicolor*). **a** The call has 11 pulses, with pulse interval less than 50 ms. Note the slow ramping of the initial 4 pulses, the amplitude plateau for the remaining 7 pulses, and the abrupt cessation of the call. All the pulses ramp on and off gradually. The second pulse (inset marked **b**) is magnified and shown in **b**. The slow rise of the call and of each pulse makes location estimation difficult for mammals

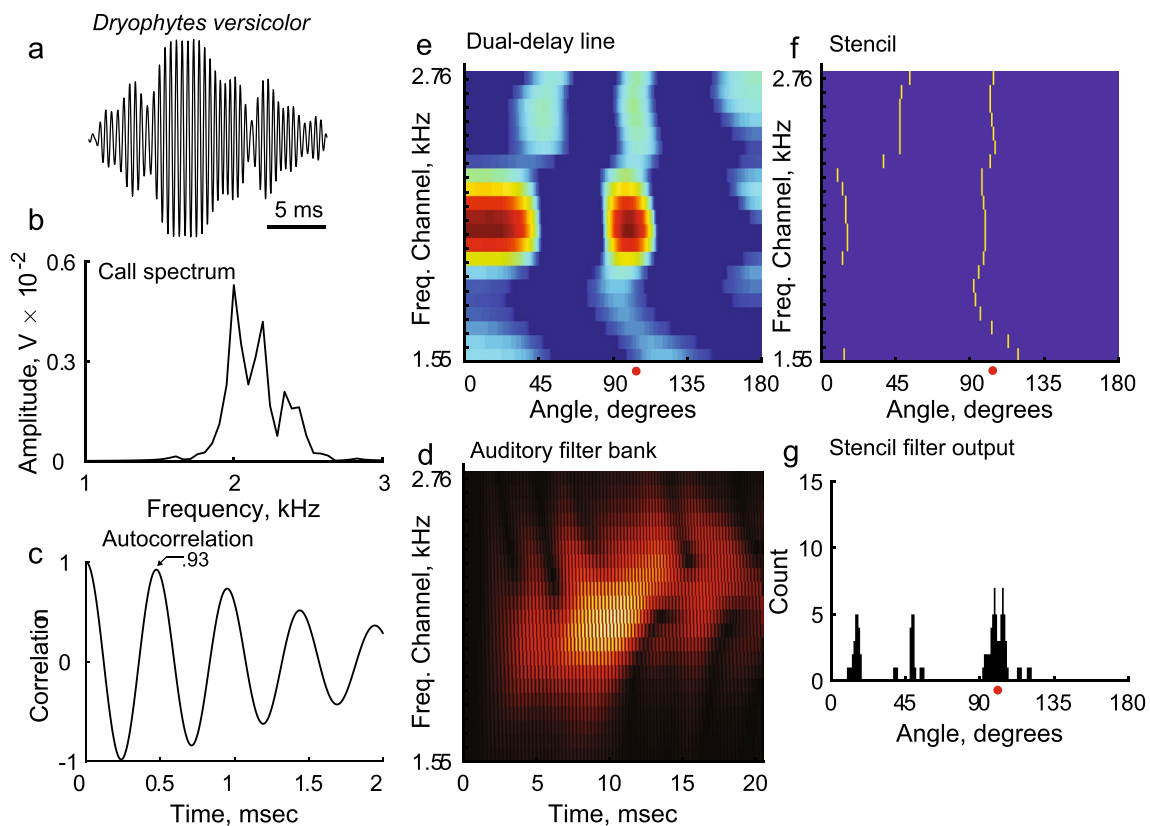


Fig. 8 Localization of a single pulse component of the gray tree frog call using mammalian ITD (interaural time difference) processing. The figure and layout are the same as Fig. 2. **a** The pulse reproduced from Fig. 6b. Noise was added to the dual delay-line inputs (20 dB

SNR). The signal is narrowband (**b**) and the second peak in the correlation function (0.93, **c**) makes the signal nearly periodic within the call and is less than 0.5 ms from the true peak. The periodicity leads to multiple false peaks (**e–g**)

true source location is at approximately 100° (i.e., about 10° from the midline, right ear leading). These reveal that the gray tree frog exploits Strategies I and II as extensively as does the cricket frog call shown earlier (Figs. 4 and 5).

Narrowband calls also defeat the population-coding strategy of the small-mammal E/I binaural model. It is important to note that population coding within the same frequency band overcomes the internal noise and quantization of discrete spiking processes, but does not overcome external environmental noise, because each neuron in the same-frequency population experiences and responds to the same noise process realization. However, noise in other, nonoverlapping frequency bands is independent, and thus cross-frequency population coding can produce gain against external environmental noise. Using experimentally recorded data from neurons across a range of frequencies, Lesica et al. (2010) used summed neuronal outputs to decode (i.e., estimate) ITDs. They showed that the summed output of about 10 neurons will reach 95% performance in a task where it is required to correctly choose one ITD out of nine ITDs (see Fig. 4e in Lesica et al. 2010). Assuming that these are independent binaural neurons, i.e., they are from non-overlapping

cochlear bands, then a simple calculation shows that ten neurons would require an unrealistically high Q to locate gray treefrogs. The call bandwidth at half-height is about 750 Hz (Fig. 8b). If there are ten non-overlapping bands (neurons) then each must have a bandwidth of 75 Hz. Assuming the center frequency of the gray treefrog call is 2 kHz, the necessary Q -factor is approximately $2000 \div 75 \approx 27$ (not to be confused with the Q value of 30 reported earlier for the spectral sharpness of the gray tree frog in Fig. 8b). The number of independent neurons at the more realistic Q value of 9 assumed in this work, is between 3 and 4, a number that may be too small to provide reliable discriminability.

The túngara frog, *Engystomops pustulosus*, call is an exception which may in fact illustrate the rule. A single call of the túngara frog (Fig. 9a, Hermans 2019), consists of a downward frequency sweep called a “whine” followed by several “chucks” (Ryan 1985; Rand and Ryan 1981). This call is referred to as a “complex” call as opposed to a “simple” call which has only a whine component and no chucks. The depicted call has 3 chucks, the first of which (Fig. 9a, marked b) is expanded and shown in Fig. 9b. We analyzed this first chuck with a dual-delay line for determining

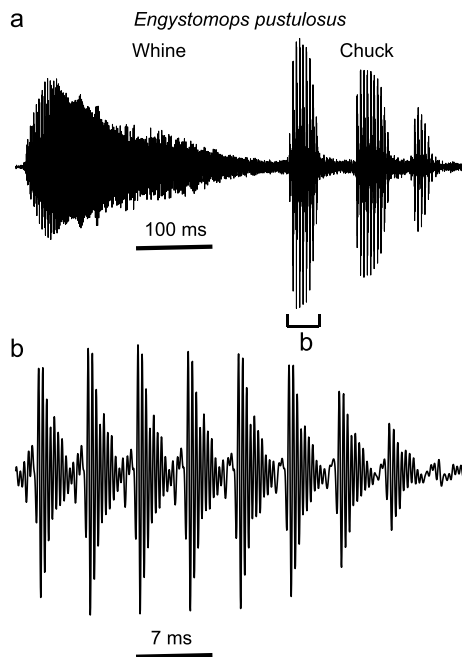


Fig. 9 The call of the túngara frog, *Engystomops pustulosus* (formerly *Physalaemus pustulosus*). **a** The “whine” and 3 “chuck” components. The whine is a downward sweeping frequency modulated signal which is not analyzed here. The first chuck (inset marked **b**) is magnified and shown in **b**. The sharp onset of each pulse making up the chuck makes this call broadband. The fringe-lipped bat *Trachops cirrhosus* preys on the túngara frog, but the chuck component of the call is more easily located than the whine (Page and Ryan 2008)

localization performance (Fig. 10). Except where noted, descriptions and methods follow those for the cricket frog and Fig. 4. Depicted are the chuck component (Fig. 10a), the spectrum of the chuck (Fig. 10b) estimated after bandpass filtering the entire call between 0.4 and 4.4 kHz, and the correlation function of the chuck (Fig. 10c). The time delay to the closest, and largest, secondary peak in the cross-correlation is 0.35 ms, with a correlation value which is 74% (−1.3 dB) of the true peak. A second channel was created by delaying the chuck by three samples. The two channels were run through identical constant-Q filter banks (Fig. 10d). Unlike Fig. 4 where the SNR was 60 dB, noise was added to produce a SNR of 20 dB (as in Fig. 5, left column). The noisy filter bank outputs were passed through a dual delay-line to produce a coincidence map (Fig. 10e), local maxima were extracted from the coincidence map to produce a “stencil” (Fig. 10f), and the stencil was summed across frequency to produce the summed coincidence map indicating the directions of acoustic sources in the environment (Fig. 10g). The true source location is at approximately 100° (i.e., about 10° from the midline, right ear leading).

It is known that the complex túngara frog call (whine plus chuck) is more easily located by predatory fringe-lipped bats *Trachops cirrhosus* (Tuttle and Ryan 1981; Page and

Ryan 2008) than the simple (whine) call alone (Page and Ryan 2008). Thus, the presence of the chuck component in the complex call presumably made the call more locatable. Closer examination of the chuck (Fig. 10a) shows that the envelope of each pulse within the chuck is asymmetric, with an abrupt onset. This leads to a spectrum of considerable bandwidth with a distinct subharmonic structure (Fig. 10b and d), and an offset correlation peak (Fig. 10c) with a much less ambiguous magnitude of 0.74 relative to that of the true peak. Furthermore, the angular dispersion of the false coincidences in the dual delay-line (Fig. 10e) and stencil (Fig. 10f) are readily apparent across the much larger bandwidth of the túngara frog chuck, producing a clear and correct location in the stencil filter output (Fig. 10g). These results support the observations made by Rand and Ryan (1981) that the chuck component of the túngara frog call is readily locatable by the fringe-lipped bat.

Why might the túngara frog forgo acoustic ambiguity? Females prefer a complex call to a simple call (Ryan 1985, Rand and Ryan 1981) and thus, there is a trade-off between sacrificing location ambiguity and call attractiveness. The spectral characteristics of the chuck are due to the túngara frog’s unusual mechanism for producing the chuck portions of its call. The complex harmonic structure of the chuck is produced by a fibrous mass connected to the frog’s vocal cords (Gridi-Papp et al., 2006). This reactive mass results in a nonlinear mixing process which generates a rich, dense, pattern of subharmonics spreading energy across a range of frequencies from about 1–4 kHz. The prominent subharmonic structure visible in the chuck spectrum in Fig. 10b is certainly suggestive of a nonlinear mixing process. This peculiar sound production mechanism increases the chuck’s bandwidth and may preclude the particular mode of stealth discussed in this section. Nevertheless, the túngara frog is known to compensate for this danger by behavioral means such as varying the relative numbers of chucks in their calls. Males generally produce simple calls (without the chuck component) when calling alone but will produce complex calls with chucks in a chorus (Page and Bernal 2006; Rand and Ryan 1981). In the next section, we will show that the very short durations of each pulse within the chuck also make it more difficult for small mammals to locate.

The whine consists of several successive down-sweeping frequency modulated (FM) harmonics with a very rapid initial frequency shift and ever slower rates of decrease as the call progresses (see Fig. 11a for the bandpass filter output as a function of frequency channel, and Fig. 11b for the frequency spectrum). The latter parts of the whine are almost periodic and rather narrow in frequency, and thus may be difficult to locate for the reasons detailed earlier in this and the previous sections. The initial rapid down-sweeping portion of the whine spans a considerable bandwidth, so it is quite interesting to observe two conveniently placed spectral

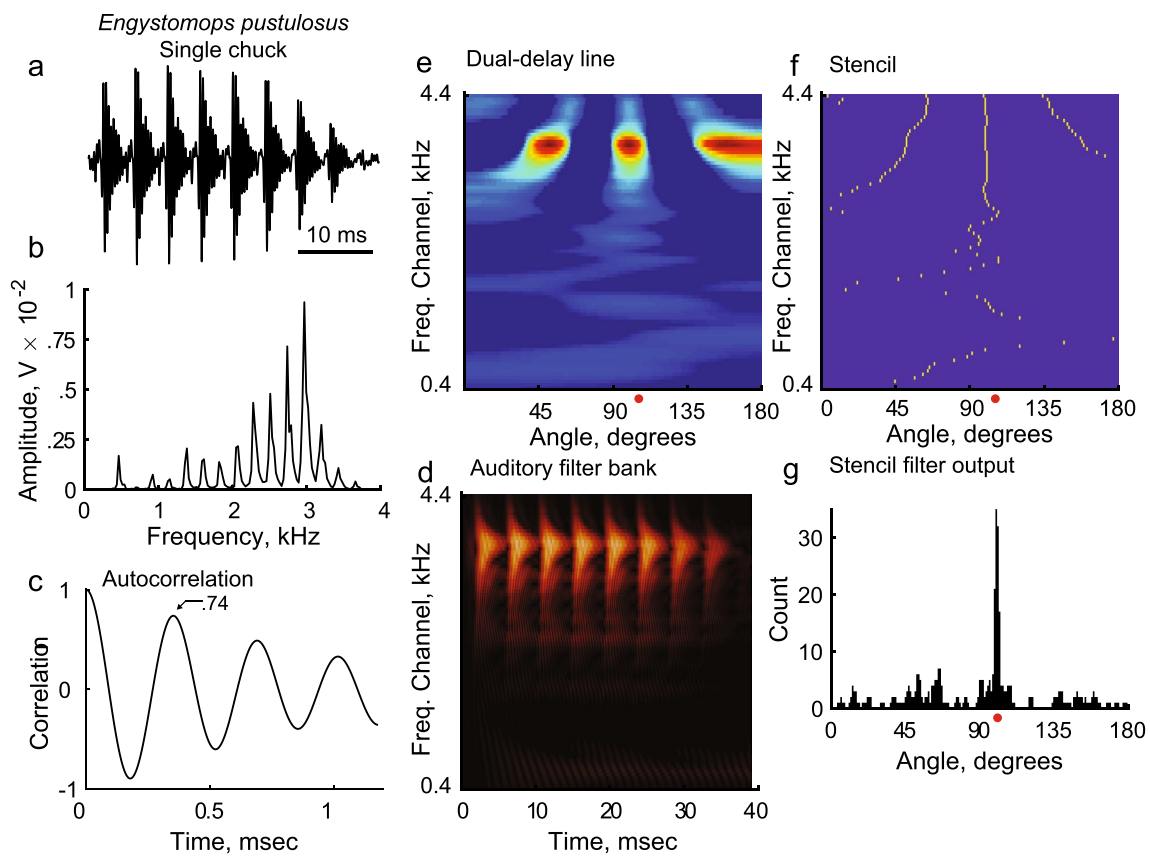


Fig. 10 Localization of a single chuck component of the túngara frog using mammalian ITD (interaural time difference) processing. The figure and layout are the same as Fig. 2. **a** The chuck component reproduced from Fig. 4b. Noise was added to the dual delay-line

inputs (20 dB SNR). Note the reduced ambiguity in the coincidence map **e** and the stencil **f**, **g**. The summed stencil filter output shows that the location estimate is unambiguous

notches just below the onset frequencies of the first and second harmonics (see white arrows, Fig. 11a, and red arrows, Fig. 11b) that restrict the bandwidth of the rapid downward sweep of the initial portion of the whine. Confining all portions of the whine to narrow frequency bands may make it harder for small mammals to locate the whine. This hypothesis could be tested by generating synthetic whines with and without these characteristics and testing them with small mammalian predators.

Strategy III: short pulses, onset detection, and the precedence effect

Each element of an extended, multi-pulse frog call presents an additional localization opportunity to the mammalian auditory system. Frogs with such calls may adopt additional strategies to mitigate this threat. We observe that many frogs calls consist of a sequence of pulses, each of short duration. For example, the duration of the túngara frog chuck and the cricket frog pulse are about 3–4 ms, the toad's pulse (not shown) is about 8 ms, and the gray tree frog pulse

ranges from 10 to 20 ms. In comparison, Brand et al. (2002) and Heffner and Heffner (1987) used 100 to 250 ms pure-tone stimuli which allow for longer integration times, and improved localization and discrimination.

The small mammal binaural E/I localization system is a triply differential mechanism that estimates small differences in time of arrival (ITD) between two closely spaced ears from a difference in responses between the left and right hemispheres, each of which is determined from small differences between ipsilateral and contralateral excitation and inhibition. Small mammal binaural localization is therefore particularly sensitive to in-band external noise. Population averaging reduces internal noise but cannot overcome external noise because each in-band neuron processes the same noise realization. However, because the noise is uncorrelated with the periodic signal, averaging over longer time intervals can produce processing gain and improve the performance. The precedence effect (Litovsky et al. 1999) seems to reset the binaural localization mechanism with each onset, and therefore deep modulations or a series of short pulses may deny the small mammal binaural

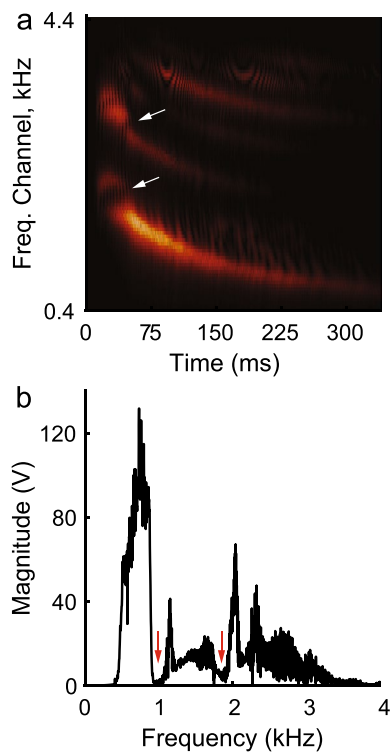


Fig. 11 Spectrogram and spectrum of the whine component of the whine-chuck call of the túngara frog depicted in Fig. 9a. **a** Spectrogram depicts four harmonic components of the downward sweeping frequency modulation of the whine. In the first and second component there is an initial rapid decrease in frequency (broadband) and a subsequent slower decrease in frequency (narrowband). Notches in the fundamental and first harmonic components (white arrows) restrict the initial high bandwidth. **a** Whine spectrum showing the placement of the two notches (red arrows)

system the longer-duration averaging intervals needed to overcome its high sensitivity to in-band noise. We note that the túngara frog's chuck has the shortest pulse duration (about 3 ms) of those we have studied, perhaps in compensation for its compromising large bandwidth.

We tested this hypothesis by generating synthetic gray tree frog pulses of durations 5, 10, 20, 50, and 100 ms and processing them with our small mammal binaural E/I model. Each synthetic pulse was created by modulating a sinusoid of frequency 2 kHz (sampling rate 80 kHz) with an envelope of the selected duration consisting of a single positive half period of a sinusoid. The signal is processed as described in Methods, with a 10 dB peak SNR at the output of the cochlear filter. 1000 noisy trials were generated for each test (depicted in Figs. 2 and 3), and the standard deviations and bias were recorded for 5 angles for each duration, as shown in Table 1. The errors shown in Table 1 range from almost random (the standard deviation of a uniform 180° distribution is 52°) for a 5 ms pulse duration, to a level comparable to the directional accuracy reported in the literature for small

Table 1 Standard deviation (SD) and bias in direction estimates of a single pulse from a synthetic gray treefrog (*Dryophytes versicolor*) call using a mammalian (excitatory/inhibitory) ITD processing model. SD and bias are in degrees

Pulse duration (ms)	5 ms SD / bias	10 ms SD / bias	20 ms SD / bias	50 ms SD / bias	100 ms SD / bias
Angle (degrees)					
-59°	44 / 20	37 / 15	30 / 10	23 / 9	18 / 9
-23°	48 / 3	38 / 3	29 / 1	18 / 3	12 / 3
0°	48 / 0	38 / 0	27 / -2	16 / 0	11 / 0
-23°	47 / -3	38 / -2	28 / -4	18 / -3	12 / -3
-59°	44 / -20	37 / -13	30 / -12	23 / -10	18 / 9

mammals tested with 100 ms noise bursts. For example, the Mongolian gerbil exhibited a 27° performance at 75% discriminability (Heffner and Heffner, 1988) and the Least Weasel (*Mustela nivalis*), a predator with a smaller head than the Mongolian gerbil exhibited a 10°–15° performance at 75% discriminability (Heffner and Heffner, 1987). However, the pulse duration of the gray treefrog ranges from 10 to 20 ms and at these durations (Table 1) the errors in directional estimates are much larger than the numbers reported for small mammals.

The standard deviation for the central angles (which are not limited on one side by the maximum realizable ITD) decreases by about $1/\sqrt{2}$ for each doubling of the pulse duration, corresponding to the expected processing gain from integrating a coherent signal against noncoherent noise. It could thus be argued that frogs gain no advantage (or loss) versus their small mammalian predators using this strategy. However, frogs call and listen primarily to attract and locate mates, so each species may optimize the parameters of their auditory system for their conspecific calls in a way that their more generalist mammalian predators cannot, thereby gaining a few dB of advantage. Individual species are known to vary greatly in their precedence preference and behavior: Marshall and Gerhardt (2010) showed a novel kind of precedence effect in a female preference for successive leading calls in a chorus rather than the first arriving pulse in *Hyla versicolor*, whereas pug nosed tree frogs (*Smilisca sila*), show no such preference (Legett et al. 2020), and *Hyla femoralis* females generally prefer trailing frog calls under conditions of partial call overlap (Merricks 2014). For relevance of the precedence effect in chorusing behavior, see Greenfield (1994a), Greenfield et al. (1997), and Legett et al. (2020).

Strategy IV: calling in chorus

Synchronized hatching, spawning, and migration are well-known strategies used by many species to reduce the risk of

predation. They temporarily overwhelm the capacity of local predators and greatly lower the risk of predation to each individual. Calling in large, dense choruses offers the same statistical protection to anurans (Ryan et al. 1981; Bradbury 1981). We focus here on signal localization in choruses.

In contrast to the linear behavior of the asymmetric pressure-gradient receiver in anurans, nonlinear operations in mammalian hearing, such as correlation and rectification, can produce local minima, false peaks, and other artifacts when signals contain multiple interfering calls. Thus, time-delay based coincidence mechanism or the binaural E/I mechanism, such as found in the mammalian auditory system, likely face greater difficulty locating an individual frog in a simultaneous chorus than does the linear asymmetric pressure-gradient receiver found in anurans. Many frog species appear to exploit this by calling simultaneously in large choruses, and by synchronizing their bouts within those choruses so as to avoid calling in isolation. It has long been argued that male frogs face a trade-off between standing out or calling in isolation to enhance their competition with other males or calling in synchrony to avoid predation; Greenfield (1994b) summarizes many selective (dis)advantages that have been hypothesized in the literature. Subsequently, Greenfield et al. (1997) argued on theoretical grounds and computer simulations that chorusing behavior is simply a byproduct of conspecific competition to jam the calls of other males, that is, that predation pressure is not necessarily required to evolve chorusing behavior. Greenfield et al. (1997) further points to the relevance of the precedence effect in inter-male call timing. A recent study by Legett et al. (2020) has experimentally measured the attraction benefits and the predation costs in the field for two species of frogs, the pug nosed tree frog and the túngara frog, and found that the pug nosed tree frog chose almost complete synchrony whereas the túngara frog chose alternating exposure in response to both the strength and direction of female selection, and the risks of predation faced by each species. Very limited field studies on this topic lend support to the general principle that species optimize their mate attraction versus predation trade-offs (Ryan et al. 1981; Legett et al. 2019), but the outcomes can be very different for each species.

The authors have constructed microphone arrays over portions of bodies of water harboring large (up to thousands of individuals) choruses of frogs and developed signal-processing techniques for isolating the locations of calling frogs and separating the calls of nearby frogs from the larger chorus (Jones and Ratnam 2009; Jones et al. 2014). This technology has enabled new insights into the chorusing behavior of frogs. Here, we focus on calling behavior within a chorus of green tree frogs, *Hyla cinerea* (reported in Jones et al. 2014).

In recordings of green tree frog choruses, individuals appear reluctant to call in isolation. Among local groups of hundreds of frogs, one individual might spontaneously begin calling once every few minutes, and it quickly ceases if no neighbor joins in a chorus. Conversely, green tree frogs are eager to join an active chorus. In our study of six neighboring frogs across 20 successive bouts at the height of a chorus (Jones et al. 2014), every frog joined in the vast majority of the bouts. The amplitude of each successive call ramps up gradually when a frog joins a bout, thereby giving other frogs time to join in and provide acoustic cover. Frogs quickly and abruptly drop out of the chorus when their neighbors cease calling.

Green tree frog choruses spontaneously organize into a three-phase calling pattern, with each frog in the chorus calling in synchrony with its brethren on one of the phases (Jones et al. 2014). There is a distinct tendency for closest neighbors to select different phases. These behaviors will have the effect of allowing a female to find a male on close approach even within a dense chorus, while jamming an interaural time-delay-based localization system with non-linear artifacts from simultaneous calls at any appreciable distance. Chief among these artifacts are false peaks in the dual delay-line coincidence due to cross-correlations between different simultaneous sources, which are unusually large due to similarity of the stereotyped calls of conspecific frogs. The stencil filter mechanism resolves these ambiguities for broadband sources, but the narrow bandwidth of most frog calls defeats that mechanism.

The gradual increase in amplitude of successive calls of green tree frogs, to make localization by mammals difficult, has a parallel in gray treefrogs. The gray tree frog's call (Fig. 7a) consists of a series of short pulses of 10–20 milliseconds in duration (Fig. 7b), repeating about every 45 milliseconds. The magnitudes of the initial pulses ramp up gradually from a very low amplitude (pulses 1 to 4, Fig. 7a), reaching a steady-state amplitude (pulses 5 to 11) but terminate abruptly (after pulse number 11). In a chorus, the gradual attack may compromise onset detection by giving the first audible arriving sound a low signal-to-noise ratio (SNR), making localization less accurate while allowing the frog to cease calling should the chorus go quiet.

Developing relative-time-delay based array signal processing algorithms to reliably locate simultaneously calling frogs in chorus was a great challenge for us; conventional algorithms failed, and we ultimately succeeded in large part only by exploiting redundancies afforded by using three widely spaced microphone arrays of five omnidirectional microphones each (Jones et al. 2014). In other work, we used an ambisonic microphone array constructed with gradient directional microphones, which are more similar to the asymmetric pressure-gradient receivers of frogs (Lockwood and Jones 2006). An ambisonic microphone array often

makes it easier to localize multiple simultaneous sources, and this was one early observation sparking the ideas for this paper.

Calling synchronously in large choruses clearly challenges the interference-sensitive small mammal binaural E/I system, as other conspecifics in the chorus are very effective in-band jammers. Despite scanning through more than an hour's worth of *Hyla cinerea* field data, we were unable to find even a single isolated pulse with SNR exceeding 10 dB. We had to look through several minutes of cricket frog and gray tree frog chorus data to find the rare instances of relatively clean calls to use in this work. Small mammalian predators face the same challenge in the wild. It may simply not be practical, to work so hard for a meal.

Discussion

Tympanic ears evolved independently in amphibians, birds, and mammals, without common ancestry (Grothe et al. 2010). Localization and directional hearing mechanisms, in particular, evolved independently in these groups. The evolution of acoustically isolated ears and binaural localization circuits in the brain in our early mammalian ancestors released them, and their descendants, from evolutionary constraints on the geometry of the head and its internal cavities required to maintain the exquisite acoustic coupling of the amphibian asymmetric pressure-difference receiver. Nonetheless, the systems are quite different, and neither is uniformly superior in all situations. This creates an opportunity to selectively play to the strengths of the pressure-difference receiver and exploit certain limitations of the mammalian system, so that frogs can acoustically hide from predators while successfully attracting conspecific mates. The inter-tympanic attenuation between the frog's ears of the pressure gradient receiver is purely a function of angle and frequency, and the shapes of the carrier waveform and envelope have no effect on this level difference over the frequency band of interest, rendering the frog's directional estimate invariant to such details. Nevertheless, it would be prudent to introduce a caveat here. In the mammalian auditory system, neurons which are ITD detectors are tuned to no more than one-half cycle of the auditory filter's center frequency (see, for example, McAlpine et al. 2001). This is the so-called π -limit (Vonderschen and Wagner 2014; McAlpine et al. 2007; Marquardt and McAlpine 2007). The π -limit may have consequences for the frog's receiver as it can introduce limitations in a pressure-gradient system. If the phase delay in the indirect tympanic pathway through the middle ear

cavity (Fig. 1a) exceeds the π -limit, then directionality will be reduced¹. This aspect requires further study. The mammalian auditory system, unlike the pressure-gradient system, estimates azimuth through central computation (whether via dual-delay lines or by a binaural E/I system). We argue that the performance of these central mechanisms is likely to be impaired by the four strategies presented here. It should be noted that mammals use their hearing for many important tasks, so selective pressure must balance performance in locating frog prey against many other behaviors. Most frogs call only to compete for and attract mates, so avoiding predation while doing so may likely exert greater selective pressure on frogs.

We have argued here that certain features of frog calls, namely nearly periodic structure at sufficiently high frequency, narrow bandwidth, short pulses, gradual onset, and calling in dense choruses, make them difficult for the mammalian ITD system to locate. But how does the mammalian auditory system process ITD? Jeffress's dual delay-line model had been accepted for decades and is considered a textbook model, but ever mounting evidence supports an excitation/inhibition (E/I), population-coding, binaural model, at least for small mammals. Based on information-theoretic principles, Harper and McAlpine (2004) have shown that optimal ITD directional processing above and below the head-width (wavelength) limit are fundamentally different, with characteristics very much like the small mammal binaural model at low frequencies, and consistent with the dual delay-line model at high frequencies (or for large heads). Rather than choose, we elected to examine both models in this paper, selecting the mammalian E/I model for a mammal with a small head (1 cm) and the dual-delay line model for a mammal with a large head (15 cm). However, unlike previous research which has addressed this problem (Harper and McAlpine 2004; McAlpine et al. 2001; Brand et al. 2002), we have focused on a specific ethological problem, namely, whether either system is able to effectively localize frog calls. We found that the dual-delay line is rendered ambiguous by periodic waveforms, and both the dual delay line and the small mammal E/I model are compromised by narrowband calls, each for somewhat different reasons. As an inherently differential system, the small mammal E/I model is particularly sensitive to noise, and we found that its performance suffers greatly in modest or low SNR conditions or high interference situations such as choruses. Frogs can maintain low SNR to defeat the small mammal binaural system by using short duration calls, or a series of short pulses (e.g., the túngara frog chuck) that presumably tricks the mammalian precedence effect into resetting the directional estimate and prevents effective integration of the call (see also Legett et al. 2020). Calling in chorus keeps the SNR low for any targeted individual and creates false correspondences in the nonlinear dual delay line.

¹ We thank one of the anonymous reviewer's for raising this interesting point.

The strategies listed here are neither mutually exclusive nor exhaustive; the first two strategies are mutually reinforcing, and many frogs use multiple strategies. The authors expect that frogs employ additional strategies that are yet to be recognized, or which are found in species we have not yet studied. As noted above, many of these difficulties are rooted in the fundamental mathematics of direction estimation based on relative time-delay and arise with digital signal processing algorithms as well. To some degree, we are arguing that the current ITD signal processing models of the mammalian auditory system (dual delay line coincidence detector; stencil filter; onset detector; or the small mammal E/I model) functionally represent the physiology with sufficient accuracy to draw conclusions, although many of the specific neurobiological mechanisms are unresolved (see Grothe et al. 2010; Feng and Schul 2007; Rose and Gooler 2007; Feng and Ratnam 2000).

Many of the ideas presented here are not new. Marler (1955, 1957) had proposed that narrow band, slow-ramping alarm calls put out by some birds are difficult to locate. This was also proposed with respect to frog advertisement calls (Rand and Ryan 1981). Alarm calls and advertisement calls are generally loud and could potentially reveal the sender's location to nearby predators. Thus, it is beneficial to make these vocalizations less readily locatable. What is perhaps new in this work is the detailed analysis and quantification of the acoustic structure of anuran calls which could potentially defeat the pressure receivers of mammals, and our complete signal processing model and implementation for the small mammal binaural E/I system.

A major drawback with this work is that we have not explicitly tested the hypotheses underlying the four strategies nor have we suggested experiments to test them. Much of the work, particularly the extensive testing of

localization in small mammals (see for example Heffner et al. 2007 and the references therein), employed 100–200 ms noise burst or tone burst stimuli. This is understandable because these earlier studies were designed to test auditory localization performance and generate comparative data. The Least weasel, in particular, is a small carnivore (Heffner and Heffner 1987) with a maximum ITD of 76 μ s, smaller than that of tested rodents including the gerbil (87 μ s). The localization accuracy of the weasel is about 10–15°, comparable with Norwegian rats (10°–13°) which seem to have the best performing localization ability among rodents (Heffner et al. 1994), although not as good as that of the dog (8°), cat (5°), or opossum (4°) (Heffner and Heffner 1987). These data are of great value and point to the importance of decoupling head-width from localization performance. What is needed from the viewpoint of anuran vocalization is a focused ethological investigation of frogs that are acoustically located by their natural predators. We have not done so here. Indeed, beyond the remarkable study with the fringe-lipped bat and túngara frog (Page and Ryan 2008; Tuttle and Ryan 1981) and the recent field study of Legett et al. (2019, 2020), we know of no other studies that quantify the difficulty of localizing anurans in the field or the laboratory as experienced by other animals, particularly mammals. Extending the problem further, it would be pertinent to ask whether avian species prey on frogs using auditory cues alone. These are ethologically relevant problems and point to a clear gap that needs to be filled.

Evolutionary biology has been criticized for generating untestable narrative explanations or hypotheses that cannot be falsified; is this paper simply another collection of “just so stories”? (Smith 2016). Fortunately, humans possess among the best directional hearing among mammals, so

Table 2 Synthetic gray treefrog calls. These calls are synthesized with interaural time disparities for binaural listening with headphones. Each file presents five instances, of which either two or three lateralize to the virtual $\pm 59^\circ$ left or right of center, respectively. The listener's task is to determine the virtual direction of each instance of

the sound source. The first filename lists the URL with the name of the folder. The remaining filenames list the name of the folder and the ‘.../’ signifies the penultimate part of the link <https://freesound.org/people/dl-jones/sounds/>

Filename	Type	Periodic Narrowband Short pulses Low onset
GrayTreefrogSynthBinaural.wav https://freesound.org/people/dl-jones/sounds/653080/	Baseline	Yes yes yes yes
GrayTreefrogSynthBinauralNarrowband.wav .../653,072/	Random carrier narrowband	No yes yes yes
GrayTreefrogSynthBinauralWideband.wav .../653,073/	Random carrier wide band	No no yes yes
GrayTreefrogSynthBinauralOnset.wav .../653,078/	Abrupt pulse onset	Yes no Yes no
GrayTreefrogSynthBinauralOffset.wav .../653,074/	Abrupt pulse offset	Yes no yes yes
GrayTreefrogSynthBinauralLongpulse.wav .../653,071/	Pulse duration tripled	Yes yes no yes

the human reader is in the rare position of at least informally testing these claims for herself. We have created a synthetic gray treefrog call closely matching the salient features of the recorded call shown in Figs. 7 and 8. We can modify this synthetic call to alter specific features that we claim affect its localizability, and then listen to a synthetic binaural presentation to evaluate that claim with our own ears.

Table 2 lists 6 sound files containing synthetic binaural gray treefrog calls embedded in 40 dB peak SNR white noise. Each file contains 5 presentations of the sound, some synthetically positioned to the left of center, and some to the right. The table lists the file names and their location on a permanent and publicly accessible site (<https://freesound.org/people/dl-jones/sounds/>), the modification made in each file, and which stealthy characteristics they remove and preserve. The files should be played through headphones as they were created for dichotic closed-field presentation. We will withhold comment on our impressions so as not to bias the reader, other than to note that to our aging ears the effect is strong in some cases and not so much in others. Of course, human low-frequency hearing is likely to exceed that of small mammals. We claim that if these sounds are not easy to locate for humans, then it is likely that they are not locatable by smaller mammals.

Acknowledgements We thank Dr. Joseph A. Dellinger for recording and providing to us the chorus data from which the gray tree frog call was extracted, and Margaret L. Jones for her invaluable assistance with the experiments that acquired the cricket frog data at the Cibolo Nature Center in Boerne, Texas. This work was supported by the University of Illinois at Urbana-Champaign (DLJ), and by a University Research Board grant (URBSASI20A5) from Ahmedabad University (RR). We thank Professor Peter Narins, our editor, for encouraging the submission of this work, and the extensive yet constructive criticisms from two anonymous reviewers. Both authors gratefully acknowledge the inspiration and guidance we received in our formative and early years in this field from our mentor, colleague, and friend, Professor Albert S. Feng. We deeply mourn his passing and dedicate this work to his memory.

Author contributions DLJ conceptualized the study, performed the analysis, and wrote the initial draft. RR provided biological background, contributed to the analysis, designed the figures, and edited the manuscript.

Declarations

Conflict of interest The authors declare no competing interests.

References

- Aaronson NL, Hartmann WM (2014) Testing, correcting, and extending the Woodworth model for interaural time difference. *J Acoust Soc Am* 135(2):817–823
- Ashida G, Carr CE (2011) Sound localization: Jeffress and beyond. *Curr Opin Neurobiol* 21(5):745–751
- Bee MA (2007) Sound source segregation in grey treefrogs: spatial release from masking by the sound of a chorus. *Anim Behav* 74(3):549–558
- 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
- Bradbury JW (1981) The evolution of leks. Natural selection and social behavior. Springer, Cham, pp 138–169
- Brand A, Behrend O, Marquardt T, McAlpine D, Grothe B (2002) Precise inhibition is essential for microsecond interaural time difference coding. *Nature* 417(6888):543–547
- Brughera AR, Stutman ER, Carney LH, Colburn HS (1996) A model with excitation and inhibition for cells in the medial superior olive. *Audit Neurosci* 2:219–233
- Carr CE, Konishi M (1990) A circuit for detection of interaural time differences in the brain stem of the barn owl. *J Neurosci* 10:3227–3246
- Cherry EC (1953) Some experiments on the recognition of speech, with one and with two ears. *J Acoust Soc Am* 25(5):975–979
- Christensen-Dalsgaard J (2005) Directional hearing in nonmammalian tetrapods. In: Popper AN, Fay RR (eds) Sound source localization. Springer, New York, pp 67–123
- Christensen-Dalsgaard J (2011) Vertebrate pressure-gradient receivers. *Hear Res* 273:37–45
- Feng AS (1980) Directional characteristics of the acoustic receiver of the leopard frog (*Rana pipiens*): a study of eighth nerve auditory responses. *J Acoust Soc Am* 68:1107–1114
- Feng AS, Ratnam R (2000) Neural basis of hearing in real-world situations. *Annu Rev Psychol* 51(1):699–725
- 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, 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
- Feng AS, Gerhardt HC, Capranica RR (1976) Sound localization behavior of the green treefrog (*Hyla cinerea*) and the barking treefrog (*H. gratiosa*). *J Comp Physiol* 107(3):241–252
- Fujita I, Konishi M (1991) The role of GABAergic inhibition in processing of interaural time difference in the owl's auditory system. *J Neurosci* 11(3):722–739
- Gabor D (1946) Theory of communication. Part 1: the analysis of information. *J IEEE-Part III* 93(26):429–441
- Gerhardt HC, Huber F (2002) Acoustic communication in insects and Anurans: common problems and diverse solutions. University of Chicago Press, Chicago
- Glasberg BR, Moore BCJ (1990) Derivation of auditory filter shapes from notched-noise data. *Hear Res* 47:103–138
- Greenfield MD (1994a) Cooperation and conflict in the evolution of signal interactions. *Annu Rev Ecol Syst* 25:97–126
- Greenfield MD (1994b) Synchronous and alternating choruses in insects and anurans: common mechanisms and diverse functions. *Am Zool* 34(6):605–615
- Greenfield MD, Tourtellot MK, Snedden WA (1997) Precedence effects and the evolution of chorusing. *Proc R Soc Lond B* 264(1386):1355–1361
- Gridi-Papp M, Rand AS, Ryan MJ (2006) Complex call production in the túngara frog. *Nature* 441(7089):38–38
- Grothe B, Sanes DH (1994) Synaptic inhibition influences the temporal coding properties of medial superior olivary neurons: an in vitro study. *J Neurosci* 14(3):1701–1709
- Grothe B, Pecka M, McAlpine D (2010) Mechanisms of sound localization in mammals. *Physiol Rev* 90:983–1012
- Harper NS, McAlpine D (2004) Optimal neural population coding of an auditory spatial cue. *Nature* 430(7000):682–686

- Harper NS, Scott BH, Semple MN, McAlpine D (2014) The neural code for auditory space depends on sound frequency and head size in an optimal manner. *PLoS One* 9(11):e108154
- Heffner RS, Heffner HE (1987) Localization of noise, use of binaural cues, and a description of the superior olivary complex in the smallest carnivore, the least weasel (*Mustela nivalis*). *Behav Neurosci* 101(5):701–708
- Heffner RS, Heffner HE (1988) Sound localization and use of binaural cues by the gerbil (*Meriones unguiculatus*). *Behav Neurosci* 102(3):422–428
- Heffner RS, Heffner HE, Kearns D, Vogel J, Koay G (1994) Sound localization in chinchillas. I: Left/right discriminations. *Hear Res* 80(2):247–257
- Heffner RS, Koay G, Heffner HE (2007) Sound-localization acuity and its relation to vision in large and small fruit-eating bats: I. Echolocating species, *Phyllostomus hastatus* and *Carollia perspicillata*. *Hear Res* 234(1–2):1–9
- Ho CC, Narins PM (2006) Directionality of the pressure-difference receiver ears in the northern leopard frog, *Rana pipiens pipiens*. *J Comp Physiol A* 192(4):417–429
- Hermans P (2019) Audio synthesis of the túngara frog call. The Digital Naturalism Conference. <https://www.dinacon.org/2019/09/11/laser-tungara-frog-synth/>. Accessed Feb 5 2022
- Jeffress LA (1948) A place theory of sound localization. *J Comp Physiol Psychol* 41:35–39
- Jones DL, Ratnam R (2009) Blind location and separation of callers in a natural chorus using a microphone array. *J Acoust Soc Am* 126(2):895–910
- Jones DL, Jones RL, Ratnam R (2014) Calling dynamics and call synchronization in a local group of unison bout callers. *J Comp Physiol A* 200(1):93–107
- Jørgensen MB (1991) Comparative studies of the biophysics of directional hearing in anurans. *J Comp Physiol A* 169(5):591–598
- Joris PX, Smith PH, Yin TC (1998) Coincidence detection in the auditory system: 50 years after Jeffress. *Neuron* 21(6):1235–1238
- Knudsen EI, Konishi M (1978) Space and frequency are represented separately in auditory midbrain of the owl. *J Neurophysiol* 41:870–884
- Koay G, Kearns D, Heffner HE, Heffner RS (1998) Passive sound-localization ability of the big brown bat (*Eptesicus fuscus*). *Hear Res* 119(1–2):37–48
- Legett HD, Page RA, Bernal XE (2019) Synchronized mating signals in a communication network: the challenge of avoiding predators while attracting mates. *Proc R Soc B* 286(1912):20191067
- Legett HD, Hemingway CT, Bernal XE (2020) Prey exploits the auditory illusions of eavesdropping predators. *Am Nat* 195(5):927–933
- Lesica NA, Lingner A, Grothe B (2010) Population coding of interaural time differences in gerbils and barn owls. *J Neurosci* 30(35):11696–11702
- Lewis ER, Narins PM (1999) The acoustic periphery of amphibians: anatomy and physiology. In: Fay RR, Popper AN (eds) *Comparative hearing: fish and amphibians*. Springer, New York, pp 101–154
- Lin WY, Feng AS (2001) Free-field unmasking response characteristics of frog auditory nerve fibers: comparison with the responses of midbrain auditory neurons. *J Comp Physiol A* 187(9):699–712
- Litovsky RY, Colburn HS, Yost WA, Guzman SJ (1999) The precedence effect. *J Acoust Soc Am* 106(4):1633–1654
- Liu C, Wheeler BC, O'Brien WD Jr, Bilger RC, Lansing CR, Feng AS (2000) Localization of multiple sound sources with two microphones. *J Acoust Soc Am* 108(4):1888–1905
- Lockwood ME, Jones DL (2006) Beamformer performance with acoustic vector sensors in air. *J Acoust Soc Am* 119(1):608–619
- Marler P (1955) Characteristics of some animal calls. *Nature* 176(4470):6–8
- Marler P (1957) Specific distinctiveness in the communication signals of birds. *Behaviour* 11(1):13–39
- Marquardt T, McAlpine D (2007) A π -limit for coding ITDs: implications for binaural models. In: Kollmeier B, Hohmann V, Klump G, Langemann U, Mauermann M, Uppenkamp S, Verhey J (eds) *Hearing-from sensory processing to perception*. Springer, Berlin, pp 407–416
- Marshall VT, Gerhardt HC (2010) A precedence effect underlies preferences for calls with leading pulses in the grey treefrog, *Hyla versicolor*. *Anim Behav* 1(1):139–145
- McAlpine D, Grothe B (2003) Sound localization and delay lines—do mammals fit the model? *Trends Neurosci* 26(7):347–350
- McAlpine D, Jiang D, Palmer AR (2001) A neural code for low-frequency sound localization in mammals. *Nat Neurosci* 4(4):396–401
- McAlpine D, Thompson S, Kriegstein KV, Marquardt T, Griffiths T, Deane-Pratt A (2007) A π -limit for coding ITDs: neural responses and the binaural display. In: Kollmeier B, Hohmann V, Klump G, Langemann U, Mauermann M, Uppenkamp S, Verhey J (eds) *Hearing-from sensory processing to perception*. Springer, Berlin, pp 399–406
- Merricks JA (2014) Coordinated communication: An analysis of signal and preference phenotypes in the genus *Hyla*. Doctoral dissertation, University of Missouri-Columbia
- Narins PM (2016) ICE on the road to auditory sensitivity reduction and sound localization in the frog. *Biol Cybern* 110(4):263–270
- Overholt EM, Rubel EW, Hyson RL (1992) A circuit for coding interaural time differences in the chick brainstem. *J Neurosci* 12(5):1698–1708
- Oxenham AJ, Shera CA (2003) Estimates of human cochlear tuning at low levels using forward and simultaneous masking. *J Assoc Res Otolaryngol* 4:541–554
- Page RA, Bernal XE (2006) Túngara frogs. *Curr Biol* 16(23):R979–R980
- Page RA, Ryan MJ (2008) The effect of signal complexity on localization performance in bats that localize frog calls. *Anim Behav* 76:761–769
- Palmer AR, Pinder AC (1984) The directionality of the frog ear described by a mechanical model. *J Theoret Biol* 110(2):205–215
- Plomp R, Mimpfen AM (1981) Effect of the orientation of the speaker's head and the azimuth of a noise source on the speech-reception threshold for sentences. *Acta Acustica United Acust* 48(5):325–328
- Rand AS, Ryan MJ (1981) The adaptive significance of a complex vocal repertoire in a neotropical frog. *Z Tierpsychol* 57(3–4):209–214
- Ratnam R, Feng AS (1998) Detection of auditory signals by frog inferior collicular neurons in the presence of spatially separated noise. *J Neurophysiol* 80(6):2848–2859
- Rose GJ, Gooler DM (2007) Function of the amphibian central auditory system. In: Narins PM, Feng AS, Fay RR, Popper AN (eds) *Hearing and sound communication in amphibians*. Springer, New York, pp 250–290
- Ryan MJ (1985) *The túngara frog: a study in sexual selection and communication*. University of Chicago Press, Chicago
- Ryan MJ, Tuttle MD, Taft LK (1981) The costs and benefits of frog chorusing behavior. *Behav Ecol Sociobiol* 8:273–278
- Saberi K, Dostal L, Sadralodabai T, Bull V, Perrott DR (1991) Free-field release from masking. *J Acoust Soc Am* 90(3):1355–1370
- Schwartz JJ, Gerhardt HC (1989) Spatially mediated release from auditory masking in an anuran amphibian. *J Comp Physiol A* 166(1):37–41
- Singheiser M, Gutfreund Y, Wagner H (2012) The representation of sound localization cues in the barn owl's inferior colliculus. *Front Neural Circuits* 6:45

- Smith RJ (2016) Explanations for adaptations, just-so stories, and limitations on evidence in evolutionary biology. *Evol Anthropol* 25:276–287. <https://doi.org/10.1002/evan.21495>
- Stern RM, Zeiberg AS, Trahiotis C (1988) Lateralization of complex binaural stimuli: a weighted-image model. *J Acoust Soc Am* 84(1):156–165
- Takahashi TT, Keller CH (1994) Representation of multiple sound sources in the owl's auditory space map. *J Neurosci* 14(8):4780–4793
- Tollin DJ, Yin TC (2003) Psychophysical investigation of an auditory spatial illusion in cats: the precedence effect. *J Neurophysiol* 90(4):2149–2162
- Tolnai S, Litovsky RY, King AJ (2014) The precedence effect and its buildup and breakdown in ferrets and humans. *J Acoust Soc Am* 135(3):1406–1418
- Tuttle MD, Ryan MJ (1981) Bat predation and the evolution of frog vocalizations in the Neotropics. *Science* 214(4521):677–678
- Tuttle MD, Taft LK, Ryan MJ (1981) Acoustical location of calling frogs by *Philander opossums*. *Biotropica* 13(3):233–234
- Vonderschen K, Wagner H (2014) Detecting interaural time differences and remodeling their representation. *Trends Neurosci* 37(5):289–300
- Woodworth RS (1938) *Experimental psychology*. Holt, Rinehart, Winston, New York
- Yin TC, Chan JC, Carney LH (1987) Effects of interaural time delays of noise stimuli on low-frequency cells in the cat's inferior colliculus. III. Evidence for cross-correlation. *J Neurophysiol* 58(3):562–583
- Young SR, Rubel EW (1983) Frequency-specific projections of individual neurons in chick brainstem auditory nuclei. *J Neurosci* 3(7):1373–1378

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.