



Research paper

Spatial release from masking in a free-field source identification task by gray treefrogs

Vivek Nityananda¹, Mark A. Bee*

Department of Ecology, Evolution and Behavior, University of Minnesota, Twin Cities, 100 Ecology, 1987 Upper Buford Circle, St. Paul, MN 55108, USA

ARTICLE INFO

Article history:

Received 25 May 2011

Received in revised form

29 December 2011

Accepted 2 January 2012

Available online 8 January 2012

ABSTRACT

Humans and other animals often communicate acoustically in noisy social groups, in which the background noise generated by other individuals can mask signals of interest. When listening to speech in the presence of speech-like noise, humans experience a release from auditory masking when target and masker are spatially separated. We investigated spatial release from masking (SRM) in a free-field call recognition task in Cope's gray treefrog (*Hyla chrysoscelis*). In this species, reproduction requires that females successfully detect, recognize, and localize a conspecific male in the noisy social environment of a breeding chorus. Using no-choice phonotaxis assays, we measured females' signal recognition thresholds in response to a target signal (an advertisement call) in the presence and absence of chorus-shaped noise. Females experienced about 3 dB of masking release, compared with a co-localized condition, when the masker was displaced 90° in azimuth from the target. The magnitude of masking release was independent of the spectral composition of the target (carriers of 1.3 kHz, 2.6 kHz, or both). Our results indicate that frogs experience a modest degree of spatial unmasking when performing a call recognition task in the free-field, and suggest that variation in signal spectral content has small effects on both source identification and spatial unmasking. We discuss these results in the context of spatial unmasking in vertebrates and call recognition in frogs.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

The “cocktail party problem” refers to the difficulty we have understanding speech when multiple people are speaking simultaneously (Bronkhorst, 2000; Cherry, 1953; McDermott, 2009). Importantly, this problem is not unique to humans and can be viewed in broad, evolutionary framework as a general problem in hearing and sound communication that we share with numerous other animals (Bee and Micheyl, 2008; Hulse, 2002). Compared to our understanding of how humans perceive speech in noisy settings, however, we know little about how nonhuman animals solve evolutionarily analogous problems (Bee and Micheyl, 2008; Hulse, 2002). Such considerations are important for understanding the mechanisms and evolution of auditory perception and vocal communication, especially in light of the evolutionary

history of hearing, which had multiple, independent origins (Webster et al., 1992). Moreover, some key features of auditory systems have even arisen independently multiple times in some lineages (Christensen-Dalsgaard and Carr, 2008; Hoy, 1992; Manley et al., 2004; Schnupp and Carr, 2009; Webster et al., 1992).

A key feature of natural soundscapes that we exploit in segregating sources is spatial separation between signals of interest and competing signals or sources of noise. Signals are more easily detected or recognized when they are separated in space from other sounds compared with co-localized conditions (Gilkey and Good, 1995; Kidd et al., 1998; Litovsky, 2005; Saberi et al., 1991; Santon, 1987; Shinn-Cunningham et al., 2005). In speech recognition tasks, for example, adults with normal hearing experience a “spatial release from masking” (SRM) of about 6–10 dB when competing speech or speech-like noise is displaced from target speech by 90° in azimuth compared with a co-localized configuration (reviewed in Bronkhorst, 2000). The purpose of the present study was to investigate SRM in a sound source identification task in frogs.

Anuran amphibians (frogs and toads) are ideally suited for studies of hearing and sound communication in noisy social settings (Feng and Schul, 2007; Narins and Zelick, 1988). Male frogs often form dense breeding choruses where they compete to attract

Abbreviations: ANOVA, analysis of variance; RMS, root mean square; SD, standard deviation; SRM, spatial release from masking; SPL, sound pressure level.

* Corresponding author. Tel.: +1 612 624 6749; fax: +1 612 624 6777.

E-mail addresses: v.nityananda@qmul.ac.uk (V. Nityananda), mbee@umn.edu (M.A. Bee).

¹ Present address: School of Biological and Chemical Sciences, Queen Mary University of London, Mile End Campus, Mile End Rd, London E1 4NS, UK.

females using advertisement calls (Gerhardt and Huber, 2002). Advertisement calls are loud (e.g., 95–110 dB peak sound pressure level at 1 m; Gerhardt, 1975), sustained ambient noise levels in active breeding choruses can be quite intense (Narins, 1982; Swanson et al., 2007), and some frog choruses can be heard from distances of up to 2 km (Arak, 1983). Within a chorus, reproductive females must be able to detect advertisement calls, localize their source, and identify the source as a male of her own species (Gerhardt and Bee, 2007). The noise in breeding choruses and the concurrent calls of nearby males can reduce signal active space (Bee, 2007; Bee and Swanson, 2007; Bee and Schwartz, 2009; Gerhardt and Klump, 1988; Wollerman, 1999), impair species recognition (Bee, 2008a; Marshall et al., 2006; Schwartz, 1987; Schwartz and Gerhardt, 1995), call type discrimination (Schwartz and Gerhardt, 1989), and source localization (Marshall et al., 2006), as well as limit the expression of female mate choice preferences (Bee, 2008b; Richardson and Lengagne, 2010; Schwartz et al., 2001; Wollerman and Wiley, 2002). In spite of these challenges, female frogs nevertheless find suitable mates in the acoustic scenes of breeding choruses. Two features of the anuran auditory system present interesting challenges to understanding how female frogs segregate sources in chorus environments.

First, frog ears function as pressure-difference receivers (Christensen-Dalsgaard, 2005, 2011; Feng and Shofner, 1981). This fact has important implications for spatial hearing in frogs. In the natural setting of a chorus, frogs may commonly encounter situations in which signals of interest and competing signals or sources of noise originate from different locations. Binaural cues for source localization are negligibly small at the *external* surfaces of the tympanic membranes given the small size of frog heads in relation to the wavelengths of sound frequencies they typically use for communication (e.g., 0.5–7 kHz) (Christensen-Dalsgaard, 2005, 2011; Gerhardt and Bee, 2007; Rheinlaender et al., 1979). The directionality of the anuran auditory periphery arises from the interaction of sound reaching both the external and internal surfaces of each tympanic membrane. Internal pathways include transmission from the contralateral tympanic membrane or from the body wall and lungs through the mouth cavity via wide Eustachian tubes (Christensen-Dalsgaard, 2005; Gerhardt and Bee, 2007). Though we have a generally good understanding of directional hearing in frogs based on behavioral and physiological studies presenting single sound sources from multiple directions (reviews in Christensen-Dalsgaard, 2005, 2011; Gerhardt and Bee, 2007; Gerhardt and Huber, 2002), we still lack detailed knowledge about how frog ears function in the presence of multiple, simultaneous sound sources (Feng and Schul, 2007). The primary objective of the present study was to provide a quantitative assessment of the extent to which their pressure-difference ears enable frogs to exploit spatial separation between signals and noise in a free-field call recognition task.

Second, amphibians are unique among vertebrates in having inner ears with two sensory papillae that encode different ranges of airborne sound frequencies. In frogs, the amphibian papilla is tonotopically organized and encodes relatively lower sound frequencies (e.g., <1.5 kHz) compared with the basilar papilla, which is broadly tuned to higher frequencies and lacks tonotopic organization (Simmons et al., 2007; Zakon and Wilczynski, 1988). In many frog species, advertisement calls have “bimodal” frequency spectra that contain separate low-frequency and high-frequency components primarily encoded by the separate papillae in the inner ear (Gerhardt and Schwartz, 2001). A secondary goal of this study was to investigate the extent to which processing of sound frequencies primarily encoded by different sensory papillae in the inner ear might contribute to a frog listener's ability to recognize calls and exploit spatial separation between signals and noise.

Recent psychophysical studies of phonotaxis behavior (approaches toward sound) with females of Cope's gray treefrog (*Hyla chrysoscelis*) suggest anurans exploit some of the same spectral (Nityananda and Bee, 2011), temporal (Vélez and Bee, 2011), and spatial (Bee, 2007, 2008a, 2010) cues as humans for perceptually organizing complex acoustic scenes. Here, we used phonotaxis experiments to test the hypothesis that spatial separation between signals and noise results in lower signal recognition thresholds (Bee and Schwartz, 2009) in a free-field source identification task. The masker was a “chorus-shaped noise” with the long-term spectrum of natural gray treefrog choruses (Fig. 1), and it was presented either co-localized with the target signal or separated by 90° in azimuth. The target signal simulated a male gray treefrog's advertisement call, which consists of a short pulse train (Fig. 1). In natural calls, each pulse contains prominent spectral energy at a fundamental frequency (and relative amplitude) of about 1.2–1.3 kHz (−6 to −10 dB) and a dominant second harmonic of about 2.4–2.6 kHz (Fig. 1). These two spectral components are primarily encoded by the amphibian and basilar papillae, respectively (Gerhardt, 2005). In the present study, we manipulated the spectral content of the target signal so that it had either the natural, “bimodal” spectrum (1.3 + 2.6 kHz) or a “unimodal” spectrum containing either just the lower (1.3 kHz) or higher (2.6 kHz) spectral peak alone (Fig. 2). Female gray treefrogs readily respond to calls with both bimodal and unimodal spectra presented at suprathreshold levels (Bee, 2010; Gerhardt, 2005; Gerhardt et al., 2007; Nityananda and Bee, 2011). This manipulation of the target signal's spectral composition allowed us to assess signal recognition thresholds and the magnitude of spatial unmasking when signals contained frequencies encoded primarily by the amphibian papilla, the basilar papilla, or both.

2. Materials and methods

2.1. Subjects

Our experiments were conducted between May 15 and July 1, 2010, with female gray treefrogs (*H. chrysoscelis*) of the western mitochondrial DNA lineage (Ptacek et al., 1994). Frogs were collected as breeding pairs in amplexus between 2130 and 0200 h from local ponds and wetlands located in the Carver Park Reserve (Carver Co., Minnesota, U.S.A.), the Crow-Hassan Park Reserve (Hennepin Co., Minnesota, U.S.A.), and the Lake Maria State Park (Wright Co., Minnesota, U.S.A.). Upon return to the laboratory, frogs were maintained at approximately 2 °C to delay egg deposition until tested. On the day of testing, subjects were placed in an incubator set to 20 °C until their body temperatures reached 20 ± 1 °C (within 30–45 min), at which time testing commenced. After testing, we reunited subjects with their chosen mates and returned them to their location of capture (usually within 48 h of collection). A total of 164 females were used as subjects in this study, which was carried out in strict accordance with recommendations in the *Guide for the Care and Use of Laboratory Animals of the National Institutes of Health*. Our experimental procedures were approved by the University of Minnesota Institutional Animal Care and Use Committee (#0809A46721).

2.2. Acoustic stimuli

2.2.1. Target signals

We conducted no-choice phonotaxis trials (Gerhardt, 1995) in which subjects were presented with an attractive target signal in the presence or absence of masking noise. We used three different target signals (Fig. 2), each consisting of a string of 32 pulses with identical gross-temporal properties that were based on average

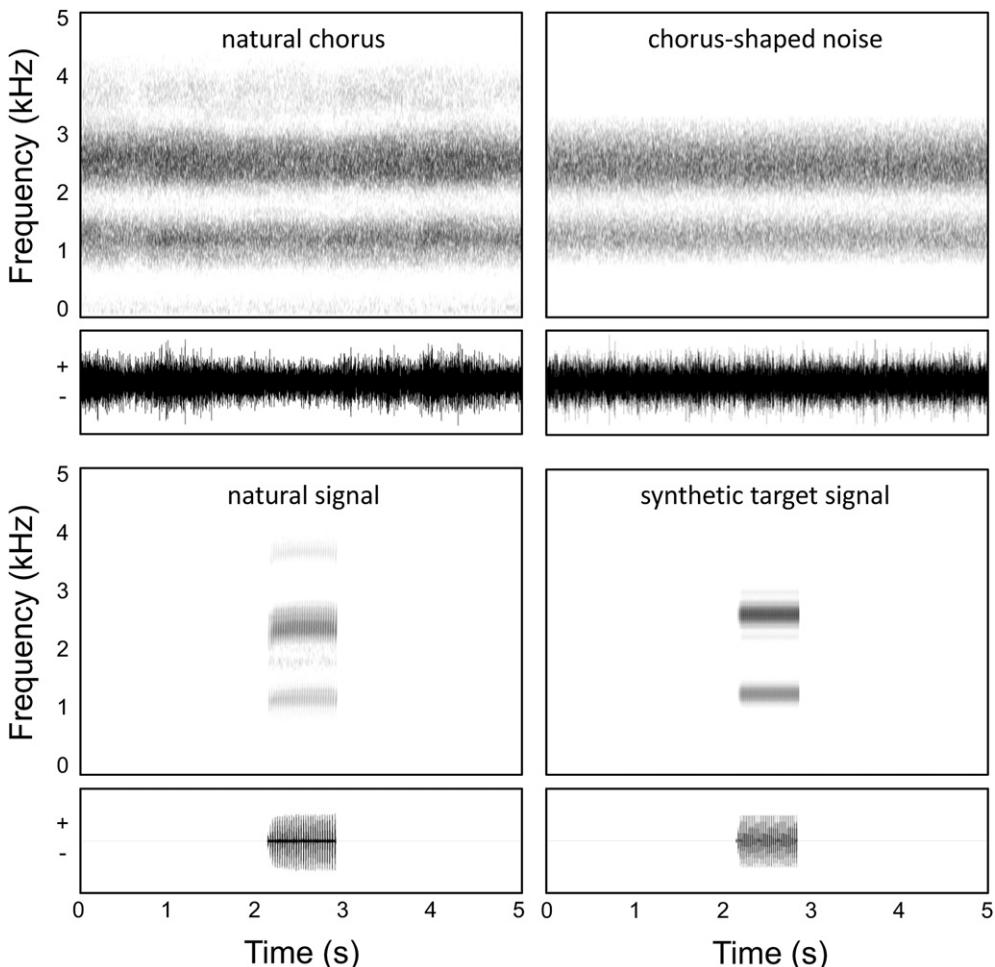


Fig. 1. Natural and artificial choruses and signals. The left column depicts spectrograms (top traces) and waveforms (bottom traces) of a segment of a natural gray treefrog chorus (top left panel) and a natural gray treefrog advertisement call (bottom left panel). The right column depicts spectrograms (top trace) and waveforms (bottom trace) of an exemplar of chorus-shaped noise used in the masking experiment (top right panel) and the bimodal target signal (bottom right panel).

values (corrected to 20 °C) of calls previously recorded in our study populations. Each pulse had a duration of 11 ms and its amplitude envelope was shaped with onsets (4 ms) and offsets (7 ms) having half-amplitude times that were 43% and 53% of the duration of the onset and offset times, respectively. Individual pulses in a target signal were separated by an 11-ms inter pulse interval so that the resulting pulse rate was 45.5 pulses s⁻¹ (50% pulse duty cycle; 693 ms signal duration). The pulses making up the three different target signals (Fig. 2) were constructed from a single sinusoid of constant frequency (1.3 kHz or 2.6 kHz for unimodal calls) or two phase-locked sinusoids (for bimodal calls) at constant frequencies of 1.3 kHz (−9 dB) and 2.6 kHz (0 dB). Target signals were shaped with a 50-ms linear onset and repeated during experiments with a period of 5 s, which approximates a natural call rate.

2.2. Chorus-shaped maskers

The maskers were conceptually based on the use of “speech-shaped noise” in studies of masked speech perception in humans. We refer to these maskers as “chorus-shaped noise” because they had the long-term spectrum of natural gray treefrog breeding choruses (Figs. 1 and 2). Between May and July, 2007–2009, we recorded fourteen different gray treefrog choruses (1.5-min recording durations) at our study sites using Marantz PMD670 digital recorders and omnidirectional Sennheiser ME62 microphones. We made recordings near the nightly peaks of calling

activity, typically between 2200 h and 0000 h. Recording microphones were placed at heights of 5 cm above ground or water level and at distances between 4 m and 10 m from the nearest calling male present in the chorus. Females in our study populations assess males from similar positions. We selected specific locations, and times of year and night, for making chorus recordings so that gray treefrogs were the only frog species heard calling. We created twenty different exemplars of chorus-shaped noise by first determining the long-term spectrum of our chorus recordings (averaged over all fourteen recordings) and then using an inverse FFT to shape the spectrum of twenty different white noises so that their long-term average spectrum matched the average chorus spectrum. Chorus-shaped noises were subsequently bandpass filtered between 0.85 kHz and 3.3 kHz.

2.3. Experimental protocol

2.3.1. Apparatus and stimulus broadcasts

Fig. 3A depicts a schematic of our testing apparatus. We conducted phonotaxis trials in a circular test arena (2-m diameter) located in a single-walled, temperature-controlled (20 ± 1 °C), hemi-anechoic sound chamber (internal L × W × H: 300 cm × 280 cm × 216 cm; Industrial Acoustics Company). With its temperature control (HVAC) unit running, the sound pressure level of the chamber's ambient noise floor ranged between 2 and

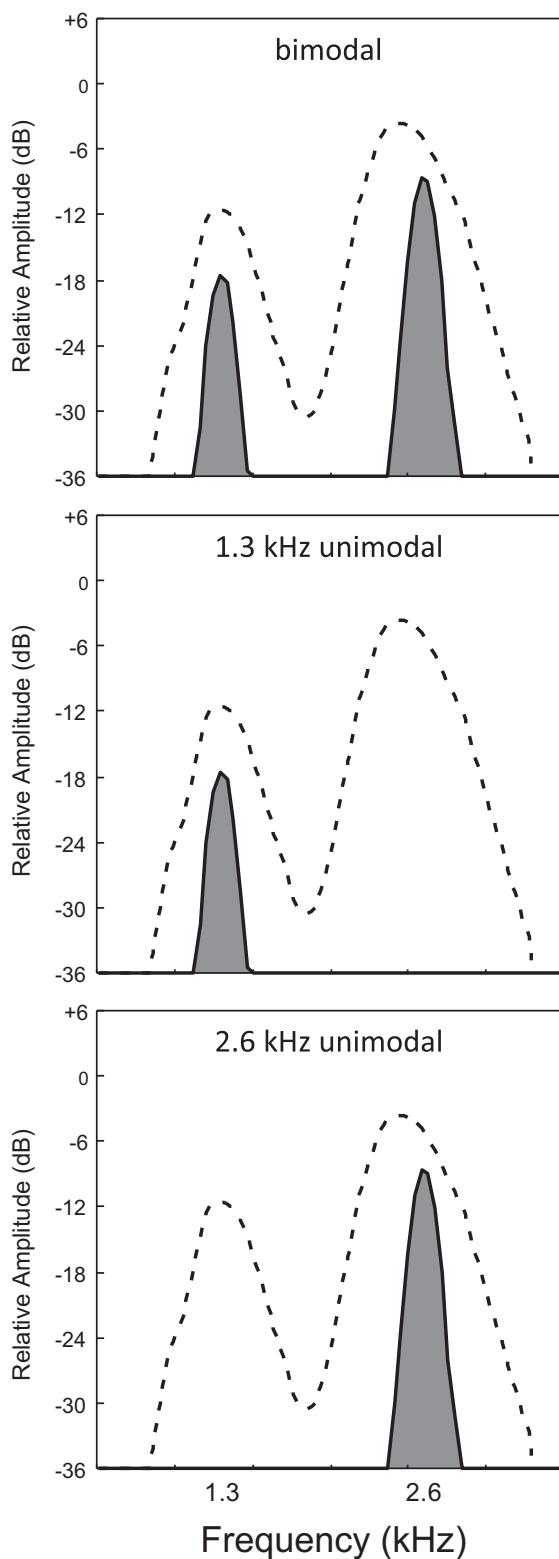


Fig. 2. Spectra of the three target signals in relation to the spectrum of chorus-shaped noise. Each plot shows the spectrum of a single target signal (shaded gray) and that of an exemplar of chorus-shaped noise used as a masker (area under the dashed line). The bimodal target signal is shown in the top panel and the two unimodal signals with carrier frequencies of 1.3 kHz and 2.6 kHz are shown in the middle and bottom panels, respectively.

12 dB SPL (re. 20 μ Pa, fast RMS, flat weighting) in the 1/3-octave bands between 500 Hz and 4000 Hz, which spans the frequency range of interest in this study. Additional details on the sound chamber have been described elsewhere (Bee and Schwartz, 2009). The wall of the test arena was 60-cm tall, constructed from hardware cloth covered in black fabric, and visually opaque but acoustically transparent. The perimeter of the arena floor was divided into 15° bins. All phonotaxis trials were conducted under IR illumination and scored in real time by two observers using a video monitor located outside the sound chamber that displayed video from an IR-sensitive camera mounted directly over the circular test arena inside the chamber (Fig. 3A). Responses were also encoded in real time as digital video files and stored to hard disk. Typically, one observer was blind to the treatment selected by the other observer. Any discrepancies between the two observers in scoring responses were resolved immediately after the trial by watching the recorded video. We have recently shown real-time scoring of phonotaxis trials to be as accurate (i.e., 100% concordance) as double-blind scoring of videos (Bee et al., in press).

We broadcast signals and maskers (20 kHz, 16-bit) from a PC located outside the sound chamber using Adobe Audition 1.5 interfaced with an M-Audio Firewire 410 soundcard. The soundcard's output was amplified (HTD 1235) and broadcast using a/d/s/ (analog and digital systems) L210 speakers that were placed on the sound chamber floor just outside the wall of the test arena, centered in one of the 15° bins, and aimed toward a subject release point at the arena's center. The frequency response of the playback setup was flat (± 3 dB). We calibrated sound levels of signals and maskers by placing the microphone of a Larson-Davis System 824 sound level meter at the approximate position of a subject's head at the central release point. The sound levels of target signals were varied across different treatments; the equivalent continuous sound level (LC_{eq}) of the chorus-shaped noise was always calibrated to be 70 dB SPL at the release point in the center of the arena. The absolute positions of speakers were systematically varied around the circular arena each testing day and between sequential tests of four to eight subjects to eliminate any possibility of a directional response bias. No such bias has ever been observed in our experimental setup.

2.3.2. Phonotaxis trials

All subjects were tested in a sequence of phonotaxis trials using an adaptive tracking procedure to measure signal recognition thresholds (see Section 2.3.3. Adaptive tracking procedure). We initiated each trial by placing a single subject in a small, circular (9-cm diameter), and acoustically transparent cage at the central release point in the test arena (Fig. 3). Subjects' initial orientations in the release cage were determined without regard for speaker positions and frogs could freely re-orient inside the cage. Each trial began with a 1-min silent period for acclimation. In conditions with masking noise, the masker started immediately following the 1-min acclimation period and continued until the end of a trial. In all conditions, broadcasts of the target signal commenced 30 s after the end of the 1-min acclimation period, and it repeated with a period of 5 s until the end of the test. All subjects were remotely released from the cage just before the fourth repetition of the target signal (i.e., ~45 s following the end of the 1-min acclimation period).

Unless indicated otherwise, we considered a "response" to have occurred when the subject's behavior met the following criteria: (i) the subject's initial contact with the arena wall was in the hemi-circle in which the target signal was broadcast; (ii) the subject contacted the arena wall in the 15° bin centered on the target signal speaker within 5 min of being released, and (iii) after touching the wall at this bin the frog remained for 30 s

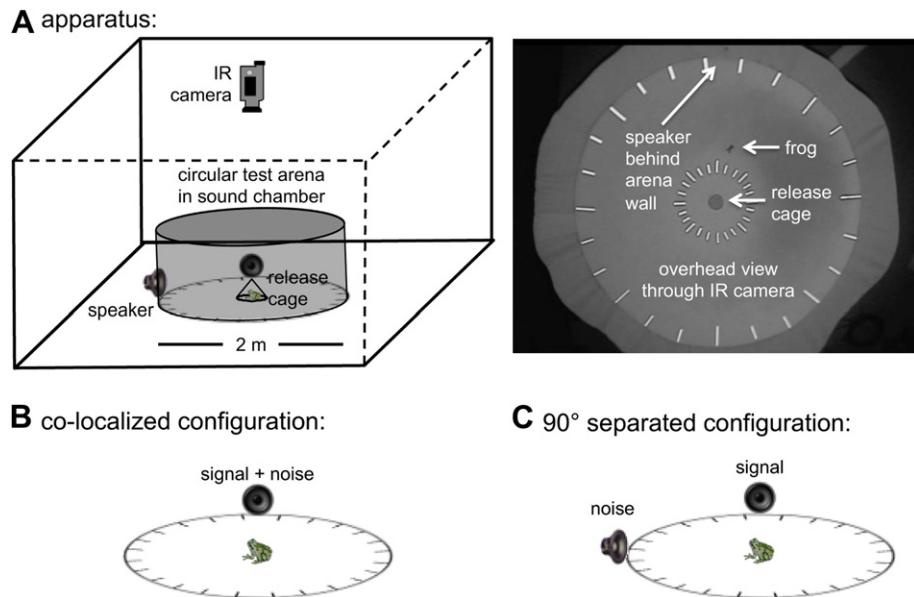


Fig. 3. Test apparatus and signal-masker configurations. A) The schematic diagram on the left illustrates the 2-m diameter circular test arena in a sound chamber and shows the positions of speakers, the central release point, and the infrared (IR) sensitive camera. On the right is shown an actual screen-shot from an IR video recording of a gray treefrog responding in a phonotaxis test. The floor or the arena was dark gray and the walls black, though the latter reflect as gray under IR illumination. B) In the co-localized configuration, signals and maskers were broadcast from the same speaker. C) In the separated conditions, signals and maskers were broadcast from two speakers positioned 90° apart around the perimeter of the circular arena. Note: frogs, speakers, and arenas are not drawn to scale in the schematic representations.

within 20 cm of the arena wall inside a bin of 30° centered on the target speaker. Subjects experienced brief “time outs” of 5–15 min in the incubator between consecutive trials in a sequence. Previous studies of treefrogs have failed to find directional biases or carry-over effects resulting from multiple tests of the same individual (Gerhardt et al., 2000).

2.3.3. Adaptive tracking procedure

Female gray treefrogs collected in amplexus exhibit high levels of motivation to respond to male advertisement calls. Phonotaxis toward a sound source in a no-choice test is interpreted as evidence that the female can (i) detect a signal, (ii) localize it, and (iii) identify the source as an appropriate mate (Beckers and Schul, 2004; Bee and Schwartz, 2009; Bush et al., 2002; Schul and Bush, 2002; Vélez and Bee, 2010, 2011). We also conducted a reference trial after any two consecutive test trials failed to elicit responses. Subjects failing to respond in any reference trial were not tested further and were excluded from statistical analyses.

2.3.3.1. *Signal recognition thresholds in quiet conditions.* We used a between-subjects design to determine signal recognition thresholds in the absence of masking noise in response to presentations of both of the unimodal signals and the bimodal signal ($N = 20$ per signal; total $N = 60$). Sequences of trials comprised both “reference trials” and “test trials,” and the total number of trials in a sequence depended on the subject’s responses. Each sequence began and ended with a reference trial, which consisted of broadcasting the bimodal target signal alone at 85 dB SPL (fast RMS, C-weighted). The reference trial simulated a single male calling at a natural amplitude (Gerhardt, 1975). Motivated

females exhibit robust phonotaxis toward target signals in this type of trial (Beckers and Schul, 2004; Bee, 2007; Bee and Swanson, 2007; Bee and Schwartz, 2009; Bush et al., 2002; Schul and Bush, 2002; Vélez and Bee, 2010, 2011). We also conducted a reference trial after any two consecutive test trials failed to elicit responses. Subjects failing to respond in any reference trial were not tested further and were excluded from statistical analyses.

In the first test trial in a sequence, we presented the designated target signal at 45 dB SPL, which is close to the signal recognition thresholds measured under quiet conditions by Bee and Schwartz (2009) using bimodal signals. In subsequent test trials in the sequence, we reduced or increased the signal level by 3 dB depending on whether the subject did or did not respond in the previous test trial, respectively. We continued either decreasing or increasing the signal level in 3-dB steps in subsequent test trials until the subject’s behavior changed (i.e., either going from response to no response, or from no response to response, between two consecutive test trials). After the subject’s behavior changed, we conducted a final test trial in which we reversed the direction of signal level change and reduced the step-size to 1.5 dB. If the subject failed to respond in this final test trial, the signal level in that trial was taken as the lower bound estimate of a signal recognition threshold and the next highest signal level previously eliciting a response was taken as the upper bound estimate. If, on the other hand, the subject responded in the final test trial, the signal level for that trial was used as the upper bound estimate of the signal recognition threshold, and the next lowest level tested in the sequence of preceding test trials was used as the lower bound estimate. We computed the signal recognition threshold as the average of the upper and lower bound using equation (1):

$$\text{Signal recognition threshold} = 10 \log_{10} \left(\frac{10^{\frac{\text{upper bound}}{10}} + 10^{\frac{\text{lower bound}}{10}}}{2} \right) \quad (1)$$

2.3.3.2. Masked signal recognition thresholds. Procedures for determining masked thresholds were similar to those described above for determining thresholds in quiet conditions with the following exceptions. First, signal recognition thresholds were measured in the presence of chorus-shaped noise for all three target signals, again using a between-subjects design. Second, we used a within-subjects design to measure two masked thresholds with signals and noise in two different spatial configurations. In the co-localized configuration (Fig. 3B), the target and masker were presented from the same speaker; in the separated configuration (Fig. 3C), the location of the target remained fixed and the location of the masker was positioned 90° laterally around the perimeter of the circular arena. We determined a subject's threshold for one configuration before determining its threshold in the second configuration. Both the order in which the two configurations were tested, and the direction the masker was displaced in the separated configuration (left or right relative to the target signal) were randomized for each subject. Finally, we conducted a "sham trial" between the first reference trial and the first test trial during which we presented the chorus-shaped noise alone, without a target signal. This trial served as a control condition to assess the possibility that subjects oriented toward the artificial chorus-shaped noises we used as maskers. Sham trials were ended when the female first touched the wall of the test arena or after 5 min had elapsed.

2.4. Statistical analyses

All of our threshold data met the requisite assumptions for using parametric statistical tests. We compared mean thresholds determined in quiet in response to the three target signals using a one-way analysis of variance (ANOVA). Our analyses of masked thresholds proceeded somewhat differently. Recall that 20 different exemplars of chorus-shaped noise were used in determining masked signal recognition thresholds. Treating each exemplar as the experimental unit of replication, we measured masked signal recognition thresholds for 1–3 subjects per exemplar in response to each of the three target signals (1.3 kHz signal, $N = 33$; 2.6 kHz signal, $N = 35$; bimodal signal, $N = 36$; total $N = 104$). Most (62%) of the 60 possible combinations of three target signals and 20 exemplars were tested with two or more subjects. When multiple subjects were tested at a particular combination of target signal and noise exemplar, we used the mean signal recognition threshold for that combination in our statistical analyses after first averaging across subjects tested with that combination (Kroodsma et al., 2001; McGregor et al., 1992). This procedure yielded a final sample size of 60 masked signal recognition thresholds. We compared these 60 thresholds using a 2 spatial configuration (within) \times 3 target signal (between) ANOVA. For all ANOVAs, we report partial η^2 as a measure of effect size, which describes the proportion of the combined effect plus error variance that can be attributed to the effect.

We conducted two separate analyses with the full dataset ($N = 104$) to assess whether subjects oriented toward the chorus-shaped masker during sham trials. First, we used circular statistics (Rayleigh tests) to test the null hypothesis that responses during the sham trials were oriented randomly with respect to position of the speaker (designated as 0°) against the alternative hypothesis that responses exhibited significant orientation (e.g., either toward or away) with respect to the speaker. Mean orientation angles close to 0° or 360° would indicate orientation toward the speaker, whereas mean angles close to 180° would indicate orientation away from the speaker. Second, we used Fisher's exact test (one-tailed) of the hypothesis that the number of subjects that first made contact with the arena wall in the 45° arc centered on

the speaker exceeded that expected based on the chance of touching the wall in any of the other seven possible 45° arcs (1/8, or 13 of 104). In our experimental setup, subjects that exhibit positive phonotaxis (e.g., in reference trials) typically make their first contact with the wall in the 15° bin in front of the speaker or one bin to either side of the speaker (Bee and Riimersma, 2008; Vélez and Bee, 2010). We used a significance criterion of $\alpha = 0.05$ for all statistical analyses.

3. Results

3.1. Thresholds in quiet conditions

In the absence of masking noise, there were no significant differences between signal recognition thresholds in response to the bimodal target signal and the two unimodal target signals ($F_{2,57} = 2.1$, $P = 0.1335$, partial $\eta^2 = 0.07$). The average (\pm SD) threshold in response to the target signal with a bimodal spectrum was 42.9 ± 5.8 dB (Fig. 4). Across subjects, threshold values in responses to this signal ranged between 29 and 55 dB. Although not statistically different from those obtained with the bimodal target signal, average thresholds in response to the 1.3 kHz ($\bar{X} = 46.6 \pm 5.0$) and 2.6 kHz ($\bar{X} = 45.5 \pm 6.9$) unimodal target signals were slightly higher. A post-hoc contrast comparing the combined mean threshold in response to both unimodal signals to that in response to the bimodal signal approached statistical significance, but the effect size associated with this comparison was still quite small ($F_{1,57} = 3.8$, $P = 0.0556$, partial $\eta^2 = 0.06$).

3.2. Masked thresholds

In a 2 spatial configuration (within) \times 3 target signal (between) ANOVA comparing masked thresholds, we found a significant effect of spatial configuration ($F_{1,57} = 28.5$, $P < 0.0001$, partial $\eta^2 = 0.33$). On average, masked thresholds were about 3 dB lower in the spatially separated configuration compared with the co-localized configuration (Fig. 4). As illustrated in Fig. 5A–C, 41 of the 60 possible combinations of 20 noise exemplars \times 3 target signals (68.3%) yielded lower thresholds in the separated configuration compared with the co-localized configuration (median difference = 4.5 dB, inter-quartile range = 3.5–11.25 dB). Eight (13.3%) exemplars yielded no difference (0 dB), and 10 (18.3%) yielded relatively higher thresholds in the co-localized configuration (median difference = 2.0 dB, inter-quartile range = 1.5–3.25 dB) (Fig. 5D–F). The two-way interaction between spatial configuration and target signal was not significant and was associated with a small effect size ($F_{2,57} = 1.23$, $P = 0.3013$, partial $\eta^2 = 0.04$), indicating that the magnitudes of threshold differences between the separated and co-localized configurations were similar for all three target signals. There was, however, a significant main effect of target signal ($F_{2,57} = 13.4$, $P < 0.0001$, partial $\eta^2 = 0.32$). Masked thresholds were lowest in response to the 1.3 kHz target signal (66.1 ± 2.6 dB), highest in response to the 2.6 kHz signal (70.4 ± 2.6 dB), and intermediate in response to the bimodal signal (68.3 ± 2.6 dB) (Fig. 4). Bonferroni post-hoc tests dissecting the main effect of target signal revealed significant differences between all pairwise comparisons (all $P < 0.0364$).

3.3. Responses in the sham trials

Of the 104 subjects tested in the sham trials, 75 made contact with the arena wall and 29 failed to do so. About half (52%) of the 29 frogs that failed to make contact with the wall actually never left the release cage, though all 104 frogs exhibited robust phonotaxis

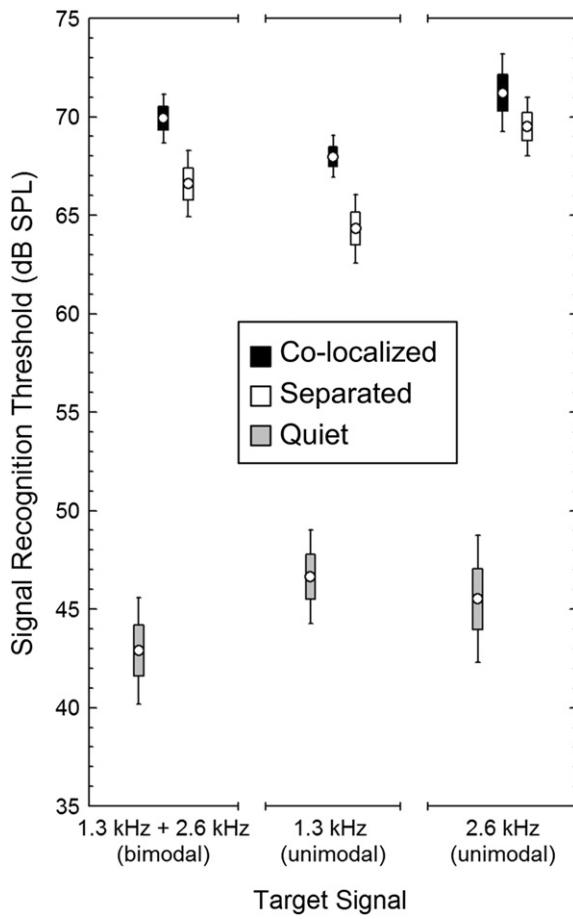


Fig. 4. Signal recognition thresholds. Depicted here are the mean (points), \pm s.e.m. (boxes), and \pm 95% confidence intervals (whiskers) for signal recognition thresholds in quiet conditions (gray) and for masked signal recognition thresholds determined in the co-localized (black) and 90° separated (white) spatial configurations.

in the reference trials. For the 75 subjects that made contact with the arena wall during the sham trial, the mean vector (μ , relative to the speaker) was 350°; however, orientation was weak (as indicated by a vector length of $r = 0.20$) and not quite statistically significant (Rayleigh test: $Z = 2.87$, $P = 0.06$, $N = 75$). For comparison, parallel analyses of responses in the first and last reference trials revealed accurate and robust orientation that was statistically significant (first: $\mu = 0.6^\circ$, $r = 0.997$, $Z = 103.33$, $P < 0.001$; last: $\mu = 0.1^\circ$, $r = 0.998$, $Z = 103.66$, $P < 0.001$; $N = 104$). Importantly, circular statistical analyses cannot take into account the behavior of the nearly 30% of subjects that never made contact with the arena wall during the sham trial. Taking the behavior of all 104 frogs into account, we found that during the sham trials, the number of subjects ($N = 20$ of 104) that made their first contact with the arena wall in the 45° arc centered on the speaker did not differ from the number expected by chance alone ($N = 13$ of 104; one-tailed Fisher's exact test: $P = 0.1724$). While female gray treefrogs exhibit positive phonotaxis toward real choruses (Christie et al., 2010; Swanson et al., 2007), three previous studies have failed to find significant orientation in response to broadcasts of artificial chorus-shaped noises (Swanson et al., 2007; Vélez and Bee, 2010, 2011). The results of the present study are consistent with these earlier reports. Based on this collective pattern of results, we conclude that orientation by female gray treefrogs toward chorus-shaped noise in the absence of a target signal is absent or very weak at best.

4. Discussion

This study yielded two main results. First, compared with a co-localized configuration, female gray treefrogs experienced, on average, about 3 dB of masking release when a chorus-shaped noise was spatially separated from a target signal by 90°. Second, both signal recognition thresholds determined in quiet and the magnitude of spatial unmasking that occurred in the separated configuration were independent of the spectral content of the signal.

4.1. Spatial unmasking in frogs

Anuran amphibians are among the most well-studied vertebrates when it comes to behavioral and physiological studies of SRM (Bee, 2007, 2008a; Lin and Feng, 2001, 2003; Ratnam and Feng, 1998; Richardson and Lengagne, 2010; Schwartz and Gerhardt, 1989, 1995). Using two-choice phonotaxis experiments, Schwartz and Gerhardt (1989) assessed the ability of female green treefrogs (*Hyla cinerea*) to detect and discriminate between attractive advertisement calls and unattractive aggressive calls in the presence of two broadband maskers. While spatial separation (90°) led to improvements in females' ability to detect the two signals, it did not improve discrimination between the two types of calls. More recent studies indicate that spatial unmasking can facilitate signal discrimination in other treefrogs and in other discrimination tasks. For example, a recent study of Cope's gray treefrogs (*H. chrysoscelis*) showed that SRM by chorus-shaped noise improved the ability of females to discriminate between conspecific and heterospecific calls (Bee, 2008a). Richardson and Lengagne (2010) reported that increased spatial separation between signals and chorus noise enhances the ability of female European treefrogs (*Hyla arborea*) to discriminate between attractive and unattractive conspecific calls.

Schwartz and Gerhardt (1989) estimated the magnitude of spatial unmasking in green treefrogs (*H. cinerea*) to be about 3 dB in a 90° separated configuration. This estimate is the same as that reported here for Cope's gray treefrogs (*H. chrysoscelis*). While small, a 3-dB release from masking is biologically relevant considering that female gray treefrogs can discriminate differences in signal levels of 2–3 dB (Bee et al., in press; Fellers, 1979; Gerhardt et al., 2000), and similarly small differences in sound levels can eliminate or reverse female preferences for calls with certain attributes (Gerhardt et al., 2000). In addition, our estimate of the magnitude of spatial unmasking, and that of Schwartz and Gerhardt (1989), are somewhat lower than that reported in a previous study of Cope's gray treefrog (Bee, 2007). In that study, Bee (2007) used no-choice phonotaxis tests to estimate population-level psychometric functions based on reaction times for responding to a target signal in the presence of a co-localized chorus-shaped noise or one separated from the target signal by 90°. Based on the observed patterns of reaction times, Bee (2007) estimated a release from masking that was at least 6 dB, but less than 12 dB. This estimate is at the high end of the range of variation in spatial release observed in the present study (see Fig. 5A–C). At present it remains unclear why Bee (2007) found more masking release compared with this study and that by Schwartz and Gerhardt (1989). Methodological differences between studies cannot be ruled out. It is also possible that differences in reaction times are more sensitive measures of spatial unmasking than differences in signal recognition thresholds.

While behavioral studies of SRM in frogs have focused on treefrogs in the genus *Hyla*, physiological studies have been conducted with the northern leopard frog (*Lithobates pipiens*, formerly *Rana pipiens*). Ratnam and Feng (1998) reported that 22% of inferior colliculus (IC) neurons had lower tone detection thresholds when a broadband noise masker was spatially

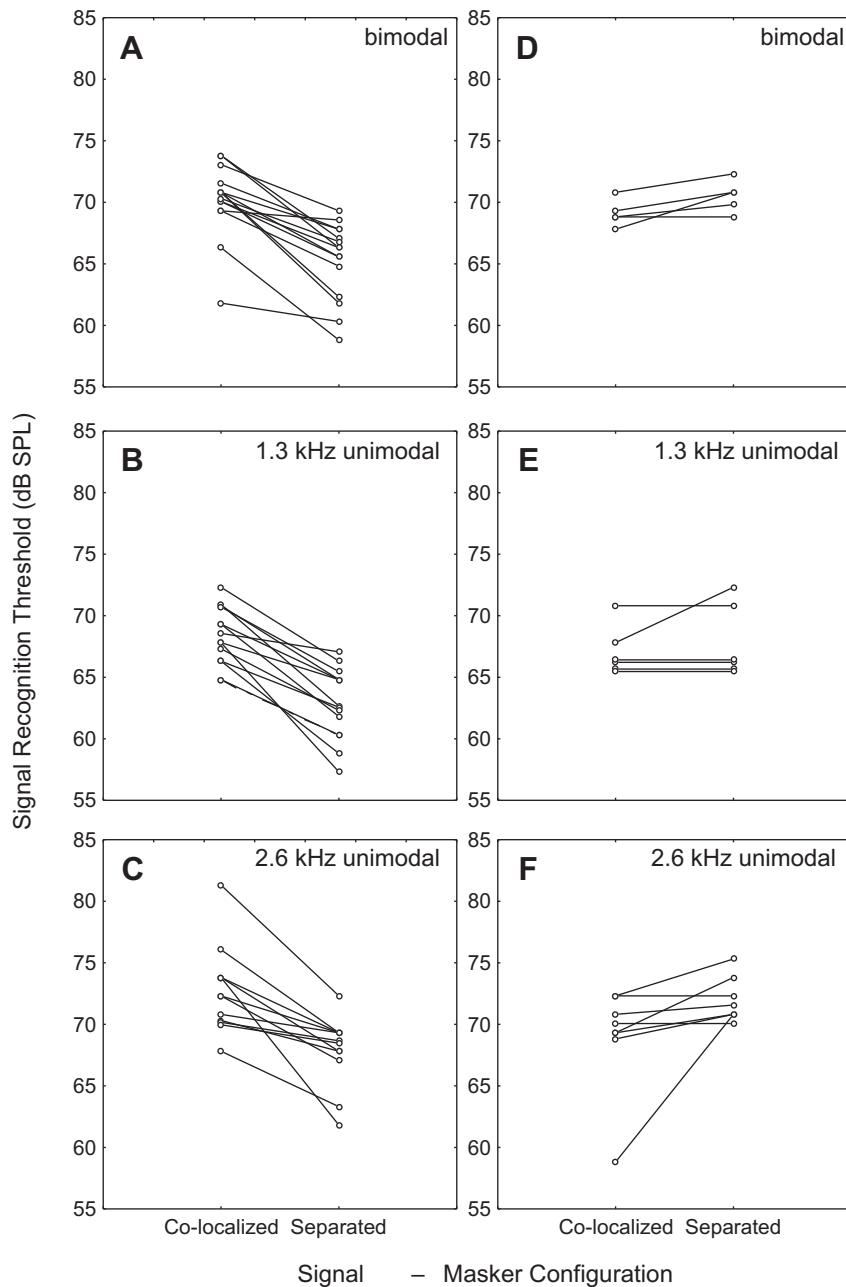


Fig. 5. Variability in masked signal recognition thresholds and spatial unmasking across noise exemplars. Plots A–C depict data for those combinations of target signal and noise exemplar that yielded relatively lower signal recognition thresholds in the separated configuration compared with the co-localized configuration. Plots D–F show equivalent data for the remaining combinations of target signal and noise exemplar for which thresholds in the separated configuration were the same as or higher than those in the co-localized condition. Each point depicts the average of 1–3 individuals tested with a given target signal and noise exemplar in the specified signal-masker configurations; pairs of points connected by a line indicate the averages of the group of individuals tested within-subjects using the same target signal and noise exemplar in both the co-localized and separated configurations.

separated (45° or 90°) from a tone bursts presented at 0° azimuth. Lin and Feng (2001) directly compared neuronal SRM in the auditory nerve and IC to investigate the contribution of central processing. The maximum SRM was about 3 dB in auditory nerve fibers and about 9 dB in the midbrain. These results confirmed that central neural processing contributed to enhancing the effect of spatial separation measured at the periphery (Lin and Feng, 2001, 2003). Lin and Feng (2003) subsequently showed that these central mechanisms involved GABAergic inhibition. Future studies integrating behavioral and physiological measures of spatial unmasking in the same species would provide for more

informative and direct comparisons of perceptual and physiological data.

4.2. Are two spectral peaks better than one?

A secondary goal of this study was to assess possible differences in the contribution to spatial unmasking and source identification of processing sound frequencies primarily encoded by the amphibian papilla, the basilar papilla, or both. There were small (2–4 dB) but significant differences in masked thresholds, with thresholds in response to the 2.6 kHz unimodal signal being

highest and those in response to the 1.3 kHz unimodal signal being the lowest. This pattern of differences was not entirely unexpected and can be explained in part by the difference in peak amplitude between the two modes of the bimodal spectrum of our chorus-shaped noises (Fig. 2). A relatively higher absolute level for the 2.6 kHz signal would be required to achieve equivalent SNRs for the two unimodal signals. Intermediate values for the bimodal target signal suggest that the addition of the 1.3 kHz peak to the 2.6 kHz, even at a relative amplitude of -9 dB, resulted in lower masked signal recognition thresholds. Similar effects also have been reported for suprathreshold presentations (Gerhardt, 2005).

Interestingly, we found no differences in SRM in tests conducted with the bimodal signal and the two unimodal signals. In fact, the two-way interaction between the two spatial configurations (co-localized versus separated) and three target signals, which tested for differences in spatial unmasking between signal types, was associated with one of the smallest effect sizes of the study (partial $\eta^2 = 0.04$). Based on these data, we conclude that spatial unmasking did not depend on whether the signal contained acoustic energy at sound frequencies primarily encoded by either the amphibian or basilar papillae, and that signals with a bimodal spectrum did not provide any additional release from masking compared with unimodal signals.

There was also little evidence that signals with two spectral peaks provided additional benefit over unimodal calls in determining signal recognition thresholds in the absence of noise. Statistical comparisons of thresholds measured in quiet conditions across the three signal types were associated with small effect sizes (partial $\eta^2 = 0.06$ – 0.07). Thresholds for responding to the bimodal signal (42.9 dB) were similar to those reported in earlier studies of this species (38 dB; Bee and Schwartz, 2009) and the closely related eastern gray treefrog, *Hyla versicolor* (37–43 dB; Beckers and Schul, 2004). Signal recognition thresholds for unimodal calls with carrier frequencies of 1.3 kHz (46.6 dB) and 2.6 kHz (45.5 dB) were similar to the average thresholds for corresponding frequencies in a pure-tone audiogram based on multi-unit recordings from the gray treefrog IC (Hillery, 1984). In that study, the audiogram exhibited a "W" shape with the highest (and equivalent) sensitivities corresponding to the two spectral peaks in the advertisement call. On the one hand, similar signal recognition thresholds for the bimodal call and the two unimodal calls are perhaps unsurprising. The gray treefrog auditory system is equally sensitive to frequencies emphasized in the two spectral peaks in the call (Hillery, 1984), and the 1.3 kHz spectral component in our bimodal signal had a naturalistic relative amplitude of -9 dB compared with the 2.6 kHz component. On the other hand, however, the observed pattern of signal recognition thresholds measured in quiet conditions is not entirely consistent with some long-standing views of call recognition in frogs.

Since the pioneering work of Capranica (1965, 1966), researchers have hypothesized that sound pattern recognition in frogs may involve "mating call detectors" that integrate sound energy across multiple regions of the spectra of vocalizations using "AND" neural computations (reviewed in Gerhardt and Huber, 2002). In support of this view, several behavioral studies have shown that call-like sounds with frequencies occurring simultaneously in two or more regions of the spectrum elicit more robust evoked calling responses from males (Capranica, 1965, 1966) and phonotaxis from females (Bee, 2010; Gerhardt, 1976, 1981, 2005; Gerhardt et al., 2007). In addition, combination sensitive neurons exist at multiple stages along the frog's ascending auditory pathway, and nonlinear facilitation can occur upon presentation of sounds comprising multiple frequency bands (Fuzessery and Feng, 1982, 1983; Mudry and Capranica, 1987a, b). Recent work in gray treefrogs has shown, however, that although females may prefer

calls with bimodal spectra, such spectra are not required for call recognition; calls with unimodal spectra presented at suprathreshold levels (e.g., 65–90 dB) elicit robust phonotaxis (Bee, 2010; Gerhardt, 2005; Gerhardt et al., 2007). However, these preferences for suprathreshold bimodal signals probably result because high sound levels excite both papillae (Gerhardt, 2005). The present study significantly extends these earlier findings by showing that unimodal signal recognition thresholds, which occur at much lower amplitudes, were largely independent of any nonlinear facilitation that may result from integrating information across frequencies. The benefit of adding the 1.3 kHz (-9 dB) component to the 2.6 kHz component was restricted to masked thresholds. While neural "AND" computations are clearly influential in determining receiver behavior across frog species, they may not always be required for call recognition.

4.3. Comparisons with other vertebrates

Comparing results from behavioral studies of free-field SRM in frogs and other vertebrates must be done with considerable caution due to large variability among studies in stimuli and experimental methods, not to mention the variability usually observed within and among subjects. In humans, for example, average SRM on the order of 0–12 dB (depending on frequency) has been reported for free-field presentations of tonal signals at 90° spatial separation compared with co-localized configurations (Gilkey and Good, 1995; Santon, 1987). In the human study perhaps most similar to ours, Litovsky (2005) reported an average free-field SRM of about 5 dB in adult listeners based on measuring speech reception thresholds in the presence of modulated speech-shaped noise in co-localized versus 90° separated configurations. Dent et al. (2009) reported an SRM of about 20–30 dB in separated (90°) versus co-localized configurations using zebra finches (*Taeniopygia guttata*) and budgerigars (*Melopsittacus undulatus*) trained to identify zebra finch songs in the presence of a chorus of similar songs or broadband noise. In an earlier study of budgerigars, Dent et al. (1997) reported an SRM of about 9 dB in a tone detection task when signal and noise were separated by 90° . Holt and Schusterman (2007) reported maximum magnitudes of spatial unmasking in an aerial tone detection task of about 19 dB and 12 dB in a harbor seal (*Phoca vitulina*) and a California sea lion (*Zalophus californianus*), respectively, when signal and masker were separated by 90° (see Turnbull, 1994, for data on underwater spatial unmasking in harbor seals). Free-field spatial unmasking has also been demonstrated in mice, *Mus musculus* (Ison and Agrawal, 1998), but that study compared co-localized and 180° separated configurations only.

On the whole, studies in other vertebrates suggest frogs may experience somewhat less SRM. At present, we do not know the reasons for these apparent differences. Among the possibilities is that species differences in body size (and consequently head size) contribute to species differences in SRM. However, differences in head size between treefrogs and small birds are probably too small to account for the large differences in spatial unmasking described above. Differences in SRM between frogs and mammals might be expected given that frog ears function as pressure-difference receivers; however, this difference alone is an insufficient explanation because the directionality of bird ears is also enhanced via a pressure-difference system (Larsen et al., 2006). Differences in experimental methodology may also explain in part, or perhaps in whole, some of the apparent differences between frogs and other vertebrates. For example, Dent et al. (1997, 2009) and Holt and Schusterman (2007) used operant conditioning paradigms; all of these studies also tested subjects in which movement of the animal in the sound field was limited. In contrast, all behavioral studies of

SRM in frogs have relied on phonotaxis as an assay, which by design requires subjects to move about in the sound field. While phonotaxis better reflects the natural listening behavior of frogs in multi-source environments, it will be important in future studies to eliminate the potential confounds introduced by mobile subjects. The reflex modification techniques used by Ison and Agrawal (1998) have also been used to study auditory processing in frogs (Hoffman and Ruppen, 1996; Simmons and Moss, 1995) and might serve as a convenient tool for investigating SRM in anuran subjects whose movements are restricted.

In contrast to our results, previous studies in other animals have found that SRM depends on the frequency of the signal. In humans, for example, the magnitude of SRM at higher sound frequencies is generally as large as, or larger than, that observed with low- and mid-frequency sounds (Gilkey and Good, 1995; Santon, 1987). Binaural interactions (e.g., interaural time differences) contribute extensively to SRM for frequencies below about 1.5 kHz, whereas the head shadow effect contributes relatively more to SRM at higher frequencies (reviewed in Bronkhorst, 2000). The situation in nonhuman animals is less clear. For example, in tests with tones centered at 1, 8, or 16 kHz in the presence of octave-wide narrowband noise centered at the same frequency, a harbor seal exhibited relatively greater spatial unmasking at 16 kHz compared with 1 and 8 kHz, which yielded similar amounts of masking release (Holt and Schusterman, 2007). A California sea lion, in contrast, exhibited the greatest spatial unmasking with a 1 kHz tone and practically no masking release with a 16 kHz tone; masking release with the 8 kHz tone was intermediate (Holt and Schusterman, 2007). In their study of mice, Ison and Agrawal (1998) presented tones at either 4 kHz or 25 kHz in the presence of octave-wide narrowband noise with the same center frequency. Spatial unmasking was observed only at the higher frequency, a finding they attributed to an effective sound shadow contralateral to the source for the 25 kHz tone but not the 4 kHz tone.

In gray treefrogs, small head size (≈ 1 cm interaural distance) probably means negligible head shadow effects at frequencies of both 1.3 kHz and 2.6 kHz (wavelengths of 26 cm and 13 cm, respectively, at 20 °C) measured at the external surfaces of the tympanic membranes (Rheinlaender et al., 1979). While anuran pressure-difference receivers are inherently directional, laser vibrometric measurements in gray treefrogs have shown the greatest directionality in a narrow frequency range between the two dominant spectral peaks in the call (Jørgensen, 1991; Jørgensen and Gerhardt, 1991). These biomechanical studies of tympanic vibrations may ultimately explain why unimodal and bimodal calls resulted in similar magnitudes of spatial unmasking in this treefrog species.

4.4. Conclusions

Results from the present study suggest SRM could contribute to solving the cocktail-party-like problem that female frogs face when choosing a mate under “real-world” listening conditions. Considering the cocktail party problem in an evolutionary framework is important because evolution is well known for creating diverse solutions to common problems (Gerhardt and Huber, 2002). Many of the basic mechanisms required for hearing in multi-source environments probably arose early in the evolution of vertebrate hearing (Fay and Popper, 2000; Popper and Fay, 1997). All vertebrates potentially share these mechanisms. However, given the evolutionary history of vertebrate auditory systems (Christensen-Dalsgaard and Carr, 2008; Manley et al., 2004; Schnupp and Carr, 2009; Webster et al., 1992), it is important to bear in mind that different vertebrate groups may also possess novel mechanisms for perceiving acoustic signals in multi-source environments that

were derived following the diversification of the major tetrapod lineages. As studies of spatial unmasking in frogs and other animals illustrate, there is certainly scope for variation in auditory processing strategies among species, further emphasizing the need for rigorous comparative studies to understand the evolution of mechanisms for hearing in noisy natural settings.

Acknowledgments

This work was supported by NIDCD DC009582. We thank, Alejandro Vélez for recordings of natural choruses and help generating chorus-shaped maskers, Mark Crawford, Madeleine Linck, John Moriarty, Ed Quinn, and Don Pereira for access to frog breeding sites, and Nate Buerkle, Brian Chicoine, Jenna Cook, Cally Espegard, Sarah Feingold, Noah Gordon, Nick Hein, Katie Heino, Johanna Henly, Shannon Hinrichs, Joe Kleinschmidt, Betsy Linehan-Skillings, James Mertz, Cathleen Nguyen, Steffen Peterson, Abby Rapacz-Van Neuren, Alejandro Vélez, and especially Sandra Tekmen for help collecting and testing frogs.

References

- Arak, A., 1983. Sexual selection by male-male competition in natterjack toad choruses. *Nature* 306, 261–262.
- Beckers, O.M., Schul, J., 2004. Phonotaxis in *Hyla versicolor* (Anura, Hylidae): the effect of absolute call amplitude. *Journal of Comparative Physiology A* 190, 869–876.
- Bee, M.A., 2007. Sound source segregation in grey treefrogs: spatial release from masking by the sound of a chorus. *Animal Behaviour* 74, 549–558.
- Bee, M.A., 2008a. Finding a mate at a cocktail party: spatial release from masking improves acoustic mate recognition in grey treefrogs. *Animal Behaviour* 75, 1781–1791.
- Bee, M.A., 2008b. Parallel female preferences for call duration in a diploid ancestor of an allotetraploid treefrog. *Animal Behaviour* 76, 845–853.
- Bee, M.A., 2010. Spectral preferences and the role of spatial coherence in simultaneous integration in gray treefrogs (*Hyla chrysoscelis*). *Journal of Comparative Psychology* 124.
- Bee, M.A., Swanson, E.M., 2007. Auditory masking of anuran advertisement calls by road traffic noise. *Animal Behaviour* 74, 1765–1776.
- Bee, M.A., Micheyl, C., 2008. The cocktail party problem: What is it? How can it be solved? And why should animal behaviorists study it? *Journal of Comparative Psychology* 122, 235–251.
- Bee, M.A., Riemersma, K.K., 2008. Does common spatial origin promote the auditory grouping of temporally separated signal elements in grey treefrogs? *Animal Behaviour* 76, 831–843.
- Bee, M.A., Schwartz, J.J., 2009. Behavioral measures of signal recognition thresholds in frogs in the presence and absence of chorus-shaped noise. *Journal of the Acoustical Society of America* 126, 2788–2801.
- Bee, M.A., Vélez, A., Forester, J.D., 2009. Sound level discrimination by Cope's gray treefrog (*Hyla chrysoscelis*) in the presence and absence of chorus-shaped noise. *Journal of the Acoustical Society of America*, in press.
- Bronkhorst, A.W., 2000. The cocktail party phenomenon: a review of research on speech intelligibility in multiple-talker conditions. *Acustica* 86, 117–128.
- Bush, S.L., Gerhardt, H.C., Schul, J., 2002. Pattern recognition and call preferences in treefrogs (Anura: Hylidae): a quantitative analysis using a no-choice paradigm. *Animal Behaviour* 63, 7–14.
- Capranica, R.R., 1965. The Evoked Vocal Response of the Bullfrog: A Study of Communication by Sound. M.I.T. Press, Cambridge, MA.
- Capranica, R.R., 1966. Vocal response of the bullfrog to natural and synthetic mating calls. *Journal of the Acoustical Society of America* 40, 1131–1139.
- Cherry, E.C., 1953. Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America* 25, 975–979.
- Christensen-Dalsgaard, J., 2005. Directional hearing in nonmammalian tetrapods. In: Popper, A.N., Fay, R.R. (Eds.), *Sound Source Localization*, vol. 25. Springer, New York, pp. 67–123.
- Christensen-Dalsgaard, J., 2011. Vertebrate pressure-gradient receivers. *Hearing Research* 273, 37–45.
- Christensen-Dalsgaard, J., Carr, C.E., 2008. Evolution of a sensory novelty: tympanic ears and the associated neural processing. *Brain Research Bulletin* 75, 365–370.
- Christie, K., Schul, J., Feng, A.S., 2010. Phonotaxis to male's calls embedded within a chorus by female gray treefrogs, *Hyla versicolor*. *Journal of Comparative Physiology A* 196, 569–579.
- Dent, M.L., Larsen, O.N., Dooling, R.J., 1997. Free-field binaural unmasking in budgerigars (*Melopsittacus undulatus*). *Behavioral Neuroscience* 111, 590–598.
- Dent, M.L., McClaine, E.M., Best, V., Ozmeral, E., Narayan, R., Gallun, F.J., Sen, K., Shinn-Cunningham, B.G., 2009. Spatial unmasking of birdsong in zebra finches (*Taeniopygia guttata*) and budgerigars (*Melopsittacus undulatus*). *Journal of Comparative Psychology* 123, 357–367.

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

Fellers, G.M., 1979. Aggression, territoriality, and mating behaviour in North American treefrogs. *Animal Behaviour* 27, 107–119.

Feng, A.S., Shofner, W.P., 1981. Peripheral basis of sound localization in anurans: acoustic properties of the frog's ear. *Hearing Research* 5, 201–216.

Feng, A.S., Schul, J., 2007. Sound processing in real-world environments. In: Narins, P.A., Feng, A.S., Fay, R.R., Popper, A.N. (Eds.), *Hearing and Sound Communication in Amphibians*. Springer, New York, pp. 323–350.

Fuzessery, Z.M., Feng, A.S., 1982. Frequency selectivity in the anuran auditory midbrain: single unit responses to single and multiple tone stimulation. *Journal of Comparative Physiology* 146, 471–484.

Fuzessery, Z.M., Feng, A.S., 1983. Mating call selectivity in the thalamus and midbrain of the leopard frog (*Rana pipiens*): single and multiunit analyses. *Journal of Comparative Physiology* 150, 333–344.

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

Gerhardt, H.C., 1976. Significance of two frequency bands in long distance vocal communication in the green treefrog. *Nature* 261, 692–694.

Gerhardt, H.C., 1981. Mating call recognition in the green treefrog (*Hyla cinerea*): importance of two frequency bands as a function of sound pressure level. *Journal of Comparative Physiology* 144, 9–16.

Gerhardt, H.C., 1995. Phonotaxis in female frogs and toads: execution and design of experiments. In: Klump, G.M., Dooling, R.J., Fay, R.R., Stebbins, W.C. (Eds.), *Methods in Comparative Psychoacoustics*. Birkhäuser Verlag, Basel, pp. 209–220.

Gerhardt, H.C., 2005. Acoustic spectral preferences in two cryptic species of grey treefrogs: implications for mate choice and sensory mechanisms. *Animal Behaviour* 70, 39–48.

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

Gerhardt, H.C., Schwartz, J.J., 2001. Auditory tuning, frequency preferences and mate choice in anurans. In: Ryan, M.J. (Ed.), *Anuran Communication*. Smithsonian Institution Press, Washington DC, pp. 73–85.

Gerhardt, H.C., Huber, F., 2002. *Acoustic Communication in Insects and Anurans: Common Problems and Diverse Solutions*. Chicago University Press, Chicago.

Gerhardt, H.C., Bee, M.A., 2007. Recognition and localization of acoustic signals. In: Narins, P.M., Feng, A.S., Fay, R.R., Popper, A.N. (Eds.), *Hearing and Sound Communication in Amphibians*, vol. 28. Springer, New York, pp. 113–146.

Gerhardt, H.C., Tanner, S.D., Corrigan, C.M., Walton, H.C., 2000. Female preference functions based on call duration in the gray tree frog (*Hyla versicolor*). *Behavioral Ecology* 11, 663–669.

Gerhardt, H.C., Martinez-Rivera, C.C., Schwartz, J.J., Marshall, V.T., Murphy, C.G., 2007. Preferences based on spectral differences in acoustic signals in four species of treefrogs (Anura: Hylidae). *Journal of Experimental Biology* 210, 2990–2998.

Gilkey, R.H., Good, M.D., 1995. Effects of frequency on free-field masking. *Human Factors* 37, 835–843.

Hillery, C.M., 1984. Seasonality of two midbrain auditory responses in the treefrog, *Hyla chrysoscelis*. *Copeia* 1984, 844–852.

Hoffman, H.S., Ruppen, F., 1996. An apparatus for the assessment of prepulse inhibition in the frog. *Behavior Research Methods Instruments & Computers* 28, 357–359.

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

Hoy, R.R., 1992. The evolution of hearing in insects as an adaptation to predation from bats. In: Webster, D.B., Fay, R.R., Popper, A.N. (Eds.), *Evolutionary Biology of Hearing*. Springer, New York, pp. 115–129.

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

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

Jørgensen, M.B., 1991. Comparative studies of the biophysics of directional hearing in anurans. *Journal of Comparative Physiology A* 169, 591–598.

Jørgensen, M.B., Gerhardt, H.C., 1991. Directional hearing in the gray tree frog *Hyla versicolor*: eardrum vibrations and phonotaxis. *Journal of Comparative Physiology A* 169, 177–183.

Kidd, G., Mason, C.R., Rohtla, T.L., Deliwala, P.S., 1998. Release from masking due to spatial separation of sources in the identification of nonspeech auditory patterns. *Journal of the Acoustical Society of America* 104, 422–431.

Kroodsma, D.E., Byers, B.E., Goodale, E., Johnson, S., Liu, W.C., 2001. Pseudoreplication in playback experiments, revisited a decade later. *Animal Behaviour* 61, 1029–1033.

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

Lin, W.Y., Feng, A.S., 2001. Free-field unmasking response characteristics of frog auditory nerve fibers: comparison with the responses of midbrain auditory neurons. *Journal of Comparative Physiology A* 187, 699–712.

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

Litovsky, R.Y., 2005. Speech intelligibility and spatial release from masking in young children. *Journal of the Acoustical Society of America* 117, 3091–3099.

Manley, G.A., Popper, A.N., Fay, R.R., 2004. Evolution of the Vertebrate Auditory System. Springer, New York.

Marshall, V.T., Schwartz, J.J., Gerhardt, H.C., 2006. Effects of heterospecific call overlap on the phonotactic behaviour of grey treefrogs. *Animal Behaviour* 72, 449–459.

McDermott, J.H., 2009. The cocktail party problem. *Current Biology* 19, R1024–R1027.

McGregor, P.K., Catchpole, C.K., Dabelsteen, T., Falls, J.B., Fusani, L., Gerhardt, H.C., Gilbert, F., Horn, A., Klump, G.M., Kroodsma, D.E., Lambrechts, M.M., McComb, K.E., Nelson, D.A., Pepperberg, I.M., Ratcliffe, L., Searcy, W.A., Weary, D.M., 1992. Design of playback experiments: the Thornbridge Hall NATO ARW consensus. In: McGregor, P.K. (Ed.), *Playback and Studies of Animal Communication*. Plenum Press, New York, pp. 1–9.

Mudry, K.M., Capranica, R.R., 1987a. Correlation between auditory thalamic area evoked responses and species-specific call characteristics II. *Hyla cinerea* (Anura: Hylidae). *Journal of Comparative Physiology A* 161, 407–416.

Mudry, K.M., Capranica, R.R., 1987b. Correlation between auditory evoked responses in the thalamus and species-specific call characteristics I. *Rana catesbeiana* (Anura, Ranidae). *Journal of Comparative Physiology A* 160, 477–489.

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

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

Nityananda, V., Bee, M.A., 2011. Finding your mate at a cocktail party: frequency separation promotes auditory stream segregation of concurrent voices in multi-species frog choruses. *PLoS ONE* 6, e21191.

Plomp, R., 1978. Auditory handicap of hearing impairment and limited benefit of hearing aids. *Journal of the Acoustical Society of America* 63, 533–549.

Plomp, R., Mimpin, A.M., 1979a. Improving the reliability of testing the speech reception threshold for sentences. *Audiology* 18, 43–52.

Plomp, R., Mimpin, A.M., 1979b. Speech reception threshold for sentences as a function of age and noise level. *Journal of the Acoustical Society of America* 66, 1333–1342.

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

Ptacek, M.B., Gerhardt, H.C., Sage, R.D., 1994. Speciation by polyploidy in treefrogs: multiple origins of the tetraploid, *Hyla versicolor*. *Evolution* 48, 898–908.

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

Rheinlaender, J., Gerhardt, H.C., Yager, D.D., Capranica, R.R., 1979. Accuracy of phonotaxis by the green treefrog (*Hyla cinerea*). *Journal of Comparative Physiology* 133, 247–255.

Richardson, C., Lengagne, T., 2010. Multiple signals and male spacing affect female preference at cocktail parties in treefrogs. *Proceedings of the Royal Society of London Series B-Biological Sciences* 277, 1247–1252.

Saberi, K., Dostal, L., Sadralodabai, T., Bull, V., Perrott, D.R., 1991. Free-field release from masking. *Journal of the Acoustical Society of America* 90, 1355–1370.

Santon, F., 1987. Detection of a pure sound in the presence of masking noise, and its dependence on the angle of incidence of the noise. *Acustica* 63, 222–228.

Schnupp, J.W.H., Carr, C.E., 2009. On hearing with more than one ear: lessons from evolution. *Nature Neuroscience* 12, 692–697.

Schul, J., Bush, S.L., 2002. Non-parallel coevolution of sender and receiver in the acoustic communication system of treefrogs. *Proceedings of the Royal Society of London Series B-Biological Sciences* 269, 1847–1852.

Schwartz, J.J., 1987. The function of call alternation in anuran amphibians: a test of three hypotheses. *Evolution* 41, 461–471.

Schwartz, J.J., Gerhardt, H.C., 1989. Spatially mediated release from auditory masking in an anuran amphibian. *Journal of Comparative Physiology A* 166, 37–41.

Schwartz, J.J., Gerhardt, H.C., 1995. Directionality of the auditory system and call pattern recognition during acoustic interference in the gray treefrog, *Hyla versicolor*. *Auditory Neuroscience* 1, 195–206.

Schwartz, J.J., Buchanan, B.W., Gerhardt, H.C., 2001. Female mate choice in the gray treefrog (*Hyla versicolor*) in three experimental environments. *Behavioral Ecology and Sociobiology* 49, 443–455.

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

Simmons, A.M., Moss, C.F., 1995. Reflex modification: a tool for assessing basic auditory function in anuran amphibians. In: Klump, G.M., Dooling, R.J., Fay, R.R., Stebbins, W.C. (Eds.), *Methods in Comparative Psychoacoustics*. Birkhäuser Verlag, Basel, pp. 197–208.

Simmons, D.D., Meenderink, S.W.F., Vassilakis, P.N., 2007. Anatomy, physiology, and function of the auditory end-organs in the frog inner ear. In: Narins, P.A., Feng, A.S., Fay, R.R., Popper, A.N. (Eds.), Hearing and Sound Communication in Amphibians, vol. 29. Springer, New York, pp. 184–220.

Swanson, E.M., Tekmen, S.M., Bee, M.A., 2007. Do female anurans exploit inadvertent social information to locate breeding aggregations? *Canadian Journal of Zoology* 85, 921–932.

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

Vélez, A., Bee, M.A., 2010. Signal recognition by frogs in the presence of temporally fluctuating chorus-shaped noise. *Behavioral Ecology and Sociobiology* 64, 1695–1709.

Vélez, A., Bee, M.A., 2011. Dip listening and the cocktail party problem in grey treefrogs: signal recognition in temporally fluctuating noise. *Animal Behaviour* 82, 1319–1327.

Webster, D.B., Fay, R.R., Popper, A.N. (Eds.), 1992. *The Evolutionary Biology of Hearing*. Springer, New York.

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

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

Zakon, H.H., Wilczynski, W., 1988. The physiology of the anuran eighth nerve. In: Fritzsch, B., Wolkowiak, W., Ryan, M.J., Wilczynski, W., Hetherington, T. (Eds.), *The Evolution of the Amphibian Auditory System*. Wiley, New York, pp. 125–155.