



Published in final edited form as:

*J Comp Psychol.* 2010 November ; 124(4): 412–424. doi:10.1037/a0020307.

## Spectral Preferences and the Role of Spatial Coherence in Simultaneous Integration in Gray Treefrogs (*Hyla chrysoscelis*)

Mark A. Bee

Department of Ecology, Evolution, and Behavior, University of Minnesota – Twin Cities

### Abstract

The perceptual analysis of acoustic scenes may often require the integration of simultaneous sounds arising from a single source. Few studies have investigated the cues that promote simultaneous integration in the context of acoustic communication in nonhuman animals. This study of Cope's gray treefrog (*Hyla chrysoscelis*) examined female preferences based on spectral features of conspecific male advertisement calls to test the hypothesis that cues related to common spatial origin promote the perceptual integration of simultaneous signal elements (harmonics). The typical advertisement call comprises two harmonically related spectral peaks near 1.1 kHz and 2.2 kHz. Subjects generally exhibited preferences for calls with two spatially coherent harmonics over alternatives with just one harmonic. When given a choice between a spatially coherent call (both harmonics originating from the same speaker) and a spatially incoherent call (each harmonic from different spatially separated speakers), subjects preferentially chose the former in the same relative proportions in which it was chosen over single-harmonic alternatives. Preferences for spatially coherent calls over spatially incoherent alternatives did not appear to result from greater difficulty localizing the spatially incoherent sources. These results are consistent with the hypothesis that spatial coherence promotes perceptual integration of simultaneous signal elements in frogs.

### Keywords

acoustic communication; auditory grouping; auditory scene analysis; grey treefrog; *Hyla chrysoscelis*

---

The perceptual organization of acoustic scenes involves parsing the complex sound pressure wave generated by multiple sources into distinct groupings that correspond to the sounds produced by different sources. This process, commonly termed “auditory scene analysis” (Bregman, 1990), involves two fundamental types of perceptual integration that play important roles in everyday hearing, such as in the perception of speech and music by human listeners. *Sequential integration* involves the perceptual binding of temporally separated sounds from one source (e.g., syllables, words, musical notes) into a coherent auditory stream that can be attended through time and segregated from other sounds.

---

Correspondence concerning this article should be addressed to Mark A. Bee, Department of Ecology, Evolution, and Behavior, University of Minnesota, 100 Ecology, 1987 Upper Buford Circle, St. Paul, Minnesota 55108. mbee@umn.edu.

This research was presented in a poster at the 30<sup>th</sup> Mid-winter Meeting of the Association for Research in Otolaryngology, Denver, Colorado, USA, February 2007, and it was supported by grants from the National Institute on Deafness and Other Communication Disorders (DC008396 and DC009582), the University of Minnesota Graduate School, and the McKnight Foundation. I thank Vince Marshall, Vivek Nityananda, Beth Pettitt, Johannes Schul, Joshua Schwartz, and Alejandro Vélez for helpful comments on earlier versions of the manuscript, John Moriarty and the Ramsey County Parks Department for access to the Tamarack Nature Center, Madeleine Linck and the Three Rivers Park District for access to the Carver Park Reserve, Ed Quinn and Mark Crawford for access to the Lake Maria State Park, and Laura Cremin, Katherine Grillaert, Melanie Harrington, Kasen Riemersma, Eli Swanson and especially Sandra Tekmen for assistance collecting and testing frogs.

*Simultaneous integration*, on the other hand, refers to the perceptual grouping of different, simultaneously occurring sound elements across the frequency spectrum (e.g., harmonics, speech formants) into a representation of a single source. Sound elements assigned to the same percept are grouped together based on their shared commonalities in a surprisingly limited number of acoustic cues, which can include onset times, fundamental frequency (i.e., harmonicity), patterns of amplitude modulation, spectral and temporal proximity, timbre, and spatial origin (Bregman, 1990; Carlyon, 2004; Darwin, 1997, 2006; Darwin & Carlyon, 1995).

Two considerations suggest that auditory scene analysis is probably fundamental to acoustic communication in many nonhuman animals (Bee & Micheyl, 2008; Feng & Ratnam, 2000; Hulse, 2002). First, like music and human speech, animal signals often comprise sequences of sound elements (e.g., pulses, chirps, syllables), and each element may comprise multiple, simultaneously produced spectral components (e.g., harmonics, formants). Second, many acoustically mediated animal behaviors take place in noisy environments that consist of multiple, simultaneously signaling individuals that may often emit signals that overlap in time and frequency (Brumm & Slabbekoorn, 2005; Langemann & Klump, 2005). Anuran amphibians (frogs and toads) represent one taxonomic group for which both of these considerations are relevant (Gerhardt & Huber, 2002). Frog calls often consist of series of repeated sound elements (e.g., multiple pulses or notes), each of which can have acoustic energy spread across the frequency spectrum (e.g., multiple harmonics). In addition, frogs often breed in dense choruses of males, and the noise generated in such social aggregations imposes constraints on communication (Narins & Zelick, 1988). Nevertheless, the calls of male frogs effectively provide an acoustic basis for species recognition, source localization, female mate choice, male-male assessment, and even individual recognition under the noisy conditions of a breeding chorus (Gerhardt & Bee, 2007; Wells & Schwartz, 2007).

While some view auditory scene analysis as a fundamental aspect of vertebrate hearing (e.g., Fay & Popper, 2000) and regard frog communication as an exemplary case of scene analysis in action (e.g., Hulse, 2002), important questions have also been raised about the extent of similarity between how frogs and other vertebrates (e.g., birds and mammals) perceptually analyze acoustic scenes (Feng & Schul, 2007). The possible mechanisms underlying simultaneous integration and across-frequency grouping represent a relevant case in point. Frogs are unique among vertebrates in having inner ears with two different, anatomically distinct sensory structures – the amphibian and basilar papillae – that are tuned to different frequency ranges of airborne sounds. Frogs generally have increased hearing sensitivity at the frequencies emphasized in conspecific vocalizations. In addition, many species have vocalizations with a bimodal frequency spectrum in which acoustic energy (e.g., harmonics) is concentrated in separate low-frequency and high-frequency regions that primarily excite the amphibian and basilar papillae, respectively. Thus, the anuran auditory periphery is often described as a “matched filter” for recognizing the calls of conspecifics and filtering out those of heterospecifics (Capranica & Moffat, 1983; Gerhardt & Schwartz, 2001; Lewis & Narins, 1999; Zakon & Wilczynski, 1988). While combination-sensitive neurons in the frog midbrain (torus semicircularis) are known to integrate simultaneous spectral information encoded by the amphibian and basilar papillae (e.g., Fuzessery & Feng, 1982), relatively little is known about the properties of simultaneous signal elements that function to promote their integration in the context of perceiving conspecific calls.

The present study investigated the spectral preferences of females of Cope’s gray treefrogs (*Hyla chrysoscelis*) and the role of spatial coherence in promoting the simultaneous integration of harmonics in male advertisement calls. Cope’s gray treefrog is the diploid member of a cryptic diploid-tetraploid species complex; the eastern gray treefrog (*Hyla versicolor*) is the tetraploid (Holloway, Cannatella, Gerhardt, & Hillis, 2006; Ptacek,

Gerhardt, & Sage, 1994). The advertisement calls of male *H. chrysoscelis* consist of a series of discrete pulses delivered at a rate of about 40–50 pulses  $s^{-1}$  (Gerhardt, 2001). Each pulse has a bimodal frequency spectrum comprising the first (1.0–1.4 kHz) and second (2.0–2.8 kHz) harmonics, with the latter produced with an amplitude that is about 6–10 dB greater than the former (Gerhardt, 2001). At low to moderate sound pressure levels, the lower and upper harmonics are primarily encoded by the amphibian and basilar papillae, respectively (Gerhardt, 2005).

The approach taken in this study was first to identify in Experiment 1 patterns of female preferences based on spectral properties (e.g., bimodal versus unimodal calls) using two-choice phonotaxis tests (Gerhardt, 1995). I then used this information in Experiment 2 to test the hypothesis that spatial coherence promotes the integration of the two harmonics into a percept of a single call having a bimodal spectrum. The basic manipulation involved giving females a choice between a spatially coherent bimodal call, for which both harmonics originated from a single speaker, and a manipulated bimodal alternative in which the two simultaneous harmonics were broadcast from two spatially separated speakers. I predicted that under conditions promoting the simultaneous integration of the two harmonics into a single auditory object, females should respond to the manipulated stimulus as if it comprised a normal bimodal spectrum. Under conditions in which there was a breakdown of simultaneous integration, I predicted that females would respond to manipulated stimuli as if they consisted of unimodal calls.

## General Method

The procedures and equipment used to collect and test females followed those described in more detail elsewhere (Bee, 2007; Bee & Swanson, 2007). Briefly, I made nightly collections of females between 5 May and 29 June, 2006, from the Tamarack Nature Center (45°6'9.5" N, 93°2'27.5" W; Ramsey County, MN), the Carver Park Reserve (44°52'48.8" N, 93°43'3.6" W; Carver County, MN), and the Lake Maria State Park (45°19'11.96"N, 93°56'37.13"W; Wright and Sherburne Counties, MN). Females were collected in amplexus, returned to the lab, and maintained at 2 °C to delay egg deposition until testing (usually the following day). One hour prior to testing, subjects were placed in an incubator until their body temperatures reached 20 ( $\pm 1$ ) °C, which was also the temperature at which subjects were tested. At the completion of testing, I released subjects at their location of capture. In total, 79 females were used as subjects for this study and together they participated in 255 individual phonotaxis tests. All procedures were approved by the University of Minnesota's Institutional Animal Care and Use Committee and complied with the National Institutes of Health (NIH) guidelines for animal use.

I tested female phonotaxis behavior under infrared (IR) illumination in a temperature-controlled, walk-in sound chamber (inside dimensions: 245 × 215 × 250 cm, L × W × H). Full details of the test chamber and playback system have been provided elsewhere (Bee, 2007; Bee & Swanson, 2007). Digital acoustic stimuli (20 kHz sampling rate, 16-bit resolution) were created using a custom-designed sound synthesis program (courtesy J. J. Schwartz). Sounds were output from a Dell Optiplex GX620 computer using Adobe Audition 1.5 and an M-Audio FireWire 410 multi-channel soundcard and broadcast through A/D/S L310 speakers using a Sonamp1230 multi-channel amplifier. The frequency response of the playback system was flat ( $\pm 2$  dB) over the frequencies used in this study. The phonotaxis responses of subjects were recorded with an overhead, IR-sensitive video camera mounted from the ceiling of the sound chamber and displayed on a video monitor located outside the sound chamber. Two people simultaneously observed behavioral responses on the monitor in real-time and one of them was responsible for recording relevant data at the completion of a test. Most responses were also digitally encoded as MPEG video files and

were available for subsequent offline analyses and could also be used to resolve immediately any inter-observer discrepancies in judgments about a subject's response. (Such instances were extremely rare.)

I conducted two-choice phonotaxis tests in a 2-m diameter circular test arena inside the sound chamber. The walls of the test arena were constructed from 60-cm high hardware cloth and covered by visually opaque, but acoustically transparent, black cloth. Playback speakers were positioned on the floor of the chamber just outside the wall of the test arena and aimed toward the center of the arena. The positions of the playback speakers around the perimeter of the test arena were systematically varied between days of testing to control for any possibility of a directional response bias in the sound chamber. At the beginning of a test, the subject was removed from the incubator and placed in a small (9-cm diameter), acoustically-transparent release cage located on the floor at the center of the test arena. After a 1-min acclimation period, I began broadcasts of the two alternating stimuli as a continuous loop, which continued over the entire duration of a test. After four repetitions of a loop, I remotely released the subject from outside the chamber. I counted a response as occurring when the subject touched the arena wall directly in front of a speaker. For a subset of tests, I also recorded the angular displacement at which the subject first made contact with the wall of the circular test area relative to the position of a designated speaker, and I measured response latency as the time between the subject's release and the time of its response.

The two stimuli in each choice test were temporally alternated by inserting equal periods of silence between the end of one alternative and the beginning of the other. Both alternatives repeated with a period of 4.5 s, which is within the range of call rates recorded in natural populations at 20°C. In several tests, one of the two alternative stimuli was an attractive *standard call* (Fig. 1a). Each pulse of the standard call was modeled after real *H. chrysoscelis* pulses and had a bimodal frequency spectrum that was constructed by adding two phase-locked sinusoids (starting phase of 0°) with frequencies (and relative amplitudes) of 1.1 kHz (-6 dB) and 2.2 kHz (0 dB). The duration of each pulse was 10 ms and its amplitude envelope was shaped with a 4.8-ms inverse exponential on-ramp and 4.8-ms exponential off-ramp. The inter-pulse interval was 10 ms so that the resulting pulse rate was 50 pulses s<sup>-1</sup>, which is within the range of pulse rates (corrected to 20°C) recorded in Minnesota populations of *H. chrysoscelis*. The standard call consisted of 45 pulses (890-ms duration) and its envelope was shaped with a 50-ms linear rise. Note that the digital versions of all stimulus alternatives described below were generated to have the same temporal properties as the standard call (e.g., starting phases of the two harmonics, call duration, call envelope, pulse duration, pulse envelope and pulse duty cycle). Subjects tested more than once were given a 5–15 min timeout in the incubator between consecutive tests. Previous studies of treefrogs have confirmed that female responses in phonotaxis tests are not biased by carry-over effects from previous tests (Gerhardt, 1981; Gerhardt, Tanner, Corrigan, & Walton, 2000). During a test, subjects were given up to 3 min to leave the release cage and up to 5 min to touch the wall in front of a speaker. Subjects that did not meet these criteria were considered nonresponsive and were not tested further.

In all tests, the standard call was presented at an overall sound pressure level (SPL) of 75 dB (re 20 µPa, fast RMS, C-weighted). I used a Brüel and Kjær Type 2250 sound level meter to calibrate the playback levels of acoustic stimuli by placing the microphone of the sound level meter at the approximate position of a female's head at the location from which females were released. A stimulus amplitude of 75 dB SPL is about 10–18 dB lower than the natural amplitude of a male calling at a distance of 1 m (range: 85–93 dB; Gerhardt, 1975). Following Gerhardt (2005), this playback level was purposefully chosen to reduce the likelihood that each harmonic would stimulate both sensory papillae in the inner ears. As

described in more detail below, the sound pressure levels of all other stimuli were adjusted in some way relative to that of the bimodal standard call.

For statistical analyses of the data, I computed standard parametric and nonparametric statistics by hand or using Statistica 7.1 (StatSoft, 2006). Estimates of effect sizes were computed by hand following Cohen (1988) or using Statistica 7.1 or G\*Power 3.1.2 (Erdfelder, Faul, & Buchner, 1996; Faul, Erdfelder, Lang, & Buchner, 2007). Circular statistics were computed using Oriana 3.0. A criterion of  $\alpha = .05$  was used to determine statistical significance.

## Experiment 1

### Methods

Experiment 1 comprised a series of six two-choice discrimination tests that were designed to investigate patterns of female preferences based on differences in the spectral content of advertisement calls. In each test, I gave females a choice between two stimuli in which I manipulated the presence/absence and relative amplitudes of the two harmonics composing each call. The purpose of these tests was twofold. First, they were designed to reveal the relative attractiveness of bimodal calls with two spatially coherent harmonics versus unimodal calls having just one harmonic, as well as the relative attractiveness of unimodal calls consisting of either just the higher or the lower harmonic. Second, these tests served as the controls for Experiment 2 (see below). As such, results from these tests established the expected proportions of females choosing calls with two spatially coherent harmonics over a stimulus having just one harmonic or the other.

In Test 1, I paired the standard call against another identical standard call to confirm that my playback setup and procedures did not produce any systematic response biases. In four additional tests, I assessed the relative attractiveness of calls with bimodal and unimodal frequency spectra. In these tests, the alternative to the bimodal standard call consisted of a unimodal call having either the 1.1 kHz or the 2.2 kHz spectral peak presented alone (Fig. 1b & 1c). I presented the unimodal alternative so that it had either the same relative amplitude as the corresponding harmonic in the standard call (Test 2, 1.1 kHz: 68 dB SPL; Test 3, 2.2 kHz: 74 dB SPL) or the same overall amplitude (Tests 4 and 5; 75 dB SPL) as the bimodal standard call. In Test 6, I paired the two unimodal calls against each other (1.1 kHz versus 2.2 kHz) and presented each with an overall amplitude of 75 dB SPL.

At least 20 females, but no more than 30 females, were tested in each of the separate discrimination tests; most subjects were tested in more than one test. Note that while sample sizes varied between tests, they were not increased to manipulate the outcomes of any particular tests of statistical significance. Rather, in those tests for which the relative proportions of the first 20 subjects that choose each alternative were not clear cut (i.e., were not close to 1:0 or 0.5:0.5), I increased the sample size by 5 or 10 additional subjects to increase my confidence that I was accurately estimating the proportions that were to serve as the expected (null) outcomes for Experiment 2. I tested the null hypothesis that 50% of females would approach each of the two alternatives using two-tailed binomial tests. The patterns of statistical significance in separate analyses using only the first 20 subjects to respond in each test were the same as those reported below. For all tests, I report the proportions ( $\hat{p}$ ) of females choosing each alternative along with the 95% binomial proportion CIs, the  $p$ -values associated with the two-tailed binomial test, and the effect size  $g$  (Cohen, 1988), defined as the deviation from the expected null probability of .50.



## Results and Discussion

Subjects exhibited no significant preferences for one alternative over the other when presented with a choice between two bimodal standard calls in Test 1 (Table 1). This result confirms that the test apparatus and procedures used in this experiment did not introduce any directional biases in responses. The bimodal standard call was generally preferred over unimodal alternatives consisting of either the 1.1 kHz or the 2.2 kHz harmonic alone. These preferences were significant when each unimodal alternative was presented at the same relative amplitude as that of the corresponding harmonic in the standard call (Tests 2 and 3; Table 1). When the unimodal alternatives were presented at the same overall amplitude (75 dB SPL) as the standard call, females still preferred the bimodal standard call over the lower-harmonic (1.1 kHz) alternative (Test 4; Table 1), and there was a nonsignificant trend ( $p = .099$ ;  $g = .17$ ) for females to prefer the standard call (2:1) over the unimodal alternative at 2.2 kHz (Test 5; Table 1). Females exhibited significant preferences for the unimodal 2.2 kHz alternative when it was paired against the unimodal 1.1 kHz alternative at equal amplitudes of 75 dB SPL (Test 6; Table 1).

The preferences based on spectral differences observed in this study exhibited several important similarities and differences when compared to those reported in earlier studies of both the tetraploid eastern gray treefrog, *Hyla versicolor* and the diploid Cope's gray treefrog, *H. chrysoscelis* (Gerhardt, 2005; Gerhardt & Doherty, 1988; Gerhardt, Martinez-Rivera, Schwartz, Marshall, & Murphy, 2007). For example, at amplitudes of 85 dB SPL, females of *H. versicolor* preferred a bimodal standard call over unimodal alternatives consisting of either the upper or lower harmonic alone (Gerhardt & Doherty, 1988; Gerhardt, et al., 2007). In parallel tests of the diploid species, *H. chrysoscelis*, females preferred a bimodal stimulus over the high-harmonic alternative, and there was a nonsignificant trend (approximately 2:1) for females to also prefer the bimodal call over the low-harmonic alternative (Gerhardt, et al., 2007). In the present study, parallel tests were conducted with the two stimulus alternatives equalized at 75 dB SPL as well as after equalizing the corresponding relative amplitudes of each harmonic. Consistent with the results of Gerhardt et al. (2007), I found that females generally preferred calls with bimodal spectra over unimodal calls (Table 1). There are, however, some interesting differences within and between the two species that are worth briefly noting.

When given a choice between two unimodal calls at 75 dB SPL, females of both *H. chrysoscelis* and *H. versicolor* from Missouri populations preferred the lower-harmonic unimodal stimulus over the higher-harmonic alternative (Gerhardt, et al., 2007). In Test 6 (Table 1), I found the opposite result: females of *H. chrysoscelis* from Minnesota populations preferred a higher-harmonic unimodal call over a lower-harmonic alternative when both stimuli were broadcast at 75 dB SPL. That this difference in preference can simply be attributed to methodological differences between studies is unlikely. A preference for higher-harmonic over lower-harmonic calls was also found in a sample of *H. chrysoscelis* females from Minnesota that was tested using the same equipment (e.g., sound chamber, computers, amplifiers, speakers) and acoustic stimuli used by Gerhardt et al. (2007) to collect data on *H. chrysoscelis* from Missouri (Bee MA and Gerhardt HC unpublished data). Instead, differences in the spectral preferences of female *H. chrysoscelis* from Missouri and Minnesota may reflect underlying evolutionary differences. Cope's gray treefrogs from my study populations belong to a western mitochondrial lineage distributed along the species' far western geographic range. The Missouri frogs tested by Gerhardt et al. (2007) were from the eastern mitochondrial lineage that is more common throughout the southeastern United States (Holloway, et al., 2006). These studies thus highlight the potential for female preferences to be evolutionarily labile within species and between closely related species.

## Experiment 2

Together, the results of Experiment 1 and those of Gerhardt et al. (2007; Gerhardt, 2005; Gerhardt & Doherty, 1988) suggest that when female gray treefrogs exhibit preferences based on differences in spectral content, these preferences are for calls with bimodal spectra over unimodal calls. Thus, as expected based on the matched filter hypothesis (Capranica & Moffat, 1983), simultaneous stimulation of both sensory organs (the amphibian and basilar papillae) results in a form of perceptual integration that renders a bimodal stimulus more attractive than a unimodal stimulus that excites primarily just one or the other organ. The aim of Experiment 2 was to test the hypothesis that spatial coherence between the two harmonics of the advertisement call promotes their simultaneous integration, and in turn, influences the preferences exhibited by females on the basis of spectral content.

### Methods

In four two-choice discrimination tests ( $N = 30$  per test), I gave subjects a choice between a spatially coherent bimodal stimulus (the standard call) and a spatially incoherent but still bimodal alternative. These two bimodal stimuli were alternated in time. To create spatial incoherence in the bimodal alternative to the standard call, I simultaneously broadcast a unimodal 1.1 kHz call and a unimodal 2.2 kHz call from two separate speakers; the two calls were phase-locked and each had the same relative amplitudes as the equivalent spectral peak in the standard call (1.1 kHz: 68 dB SPL; 2.2 kHz: 74 dB SPL; Fig. 1b–c). In separate tests, the two speakers were separated by an angle ( $\theta$ ) that equaled  $7.5^\circ$ ,  $15^\circ$ ,  $30^\circ$ , or  $60^\circ$  (Fig. 2a). For the spatially incoherent alternatives, I situated the two spatially separated speakers so that their midpoint along the arc of the arena wall was directly opposite ( $180^\circ$ ) from the single speaker broadcasting the standard call (Fig. 2a). The relative positions of the two speakers broadcasting the spatially incoherent bimodal stimulus were systematically varied between testing days. There was little evidence to suggest that stimulus artifacts were introduced by simultaneously broadcasting two harmonically related sounds from different speakers in the spatially incoherent conditions (see Supplemental Material).

Based on the results of Experiment 1 (Table 1), I made the following predictions for Experiment 2. First, I predicted that subjects would choose the bimodal standard call and the spatially incoherent bimodal alternative in equal proportions if they perceptually integrated the two spatially separated harmonics of the latter into a percept of an attractive bimodal call. Alternatively, I predicted that subjects would prefer the bimodal standard call if the physically separated harmonics composing the spatially incoherent bimodal call were not perceptually integrated together. Based on the outcome of Experiment 1, I tested these two predictions using one-tailed binomial tests of the null hypothesis that 50% of subjects would choose the standard call against the alternative hypothesis that greater than 50% of subjects would choose the spatially coherent standard call. I counted a response to the spatially incoherent bimodal alternative as occurring when a female touched the arena wall in front of either speaker broadcasting one of the two spatially separated harmonics. For these tests, I report the proportions ( $\hat{p}$ ) of females choosing each alternative along with the 95% binomial proportion CIs [in brackets], the results of the one-tailed binomial test, and the effect size  $g$ .

I additionally predicted that if subjects failed to integrate the spatially separated harmonics together, then the proportions of subjects preferring the bimodal standard call over each harmonic in the spatially incoherent alternative should be similar in magnitude to those that preferred the bimodal standard call over the corresponding unimodal alternatives in Experiment 1. Recall that in Test 2 of Experiment 1 (Table 1), 100% of subjects ( $\hat{p} = 1.0$ , 95% CI [.83, 1.0],  $N = 20$ ) chose the bimodal standard call over the unimodal 1.1 kHz alternative, whereas subjects in Test 3 (Table 1) chose the standard call over the unimodal 2.2 kHz alternative in a 5.25:1 ratio, or in proportions of  $\hat{p} = .84$  (95% CI [.63, .95],  $N = 25$ )

and  $\hat{p} = .16$  (95% CI [.05, .36],  $N = 25$ ), respectively. When given a choice in Test 6 (Table 1) between the two unimodal calls, one with a spectral peak at 1.1 kHz and the other with a peak at 2.2 kHz, none of subjects chose the lower-harmonic (1.1 kHz) alternative ( $\hat{p} = .00$ , 95% CI [0, .17],  $N = 20$ ). Therefore, if spatial coherence promoted simultaneous integration, a reasonable expectation based on these results was that the subjects tested in Experiment 2 should approach the single speaker broadcasting the bimodal standard call and the two speakers broadcasting the spatially separated unimodal 1.1 kHz and 2.2 kHz calls in expected proportions of 0.84, 0.0, and 0.16, respectively (Fig. 2b–e). I used  $z$  tests of two independent proportions to evaluate the hypothesis that the observed proportions of subjects that responded to each speaker did not differ from these null expected proportions. I also calculated Cohen's (1988)  $h$  as a measure of effect size for these tests. Note that the separation between the bimodal stimulus and the unimodal alternatives was  $180^\circ$  in Experiment 1, but varied between  $150^\circ$  and  $172.5^\circ$  as a function of  $\theta$  in Experiment 2. These analyses thus assume that this difference had no effect on the patterns of female preferences.

The use of phonotaxis as a behavioral assay to measure auditory grouping based on spatial cues is potentially confounded by possible differences in source localizability (Bee & Riemersma, 2008). In other words, subjects in Experiment 2 might have avoided responding to a spatially incoherent bimodal stimulus not because of a failure to integrate its separate elements together (i.e., a breakdown of grouping), but because the grouped percept could not be assigned to a definite location (i.e., breakdown of source localization). I evaluated this possibility in two ways.

First, for each magnitude of spatial incoherence ( $\theta$ ), I used one-way analysis of variance (ANOVA) to simultaneously compare  $\log_{10}$ -transformed latencies in response to the spatially coherent and incoherent stimuli of this experiment with those from Test 1 of Experiment 1, in which females were given a choice between two bimodal standard calls. The  $\log_{10}$  transformation was used to bring the data in line with the normality assumption of ANOVA. For these ANOVAs, I used a sample size of  $N = 19$  for Experiment 1 after removing one subject also tested in Experiment 2. I report partial  $\eta^2$  as a measure of effect size for these analyses. Sample sizes for these analyses reflect the total number of subjects for which response latencies were available. The general expectation was that difficulty in source localization should be reflected in relatively longer latencies for responses to one of the physically separated harmonics composing the spatially incoherent bimodal call (Bee & Riemersma, 2008; Feng, Gerhardt, & Capranica, 1976).

Second, I used  $\chi^2$  tests (Batschelet, 1981) to make between-groups comparisons of the angles at which subjects touched the arena wall, relative to the speaker they eventually chose. In all analyses, the control group comprised the subjects given a choice between two spatially coherent bimodal calls in Test 1 of Experiment 1 ( $N = 19$  after removing one subject also tested in Experiment 2). In separate analyses, the treatment groups consisted of subjects that chose either of the speakers broadcasting one of the spatially separated harmonics at the separate levels of  $\theta$ . These analyses were performed after designating the position of the chosen speaker as  $0^\circ$  and the direction of the other spatially separated speaker as being toward  $90^\circ$ . These  $\chi^2$  analyses were designed to determine whether the two groups being compared had different means or distributions of response angles. The expectation was that greater difficulty localizing a spatially incoherent source would be reflected in one or both of these parameters.

## Results and Discussion

Females generally preferred the spatially coherent, bimodal standard call when it was paired against a spatially incoherent bimodal alternative. Across all levels of spatial separation, between 63–73% of females chose the single speaker broadcasting the standard call over one



of the other two spatially separated speakers broadcasting the spatially incoherent bimodal alternative (Fig. 2b–e). These preferences were significant at the  $7.5^\circ$  ( $\hat{p} = .73$ , 95% CI [.57, 1.0],  $p = .008$ ,  $g = .23$ ),  $30^\circ$  ( $\hat{p} = .70$  [.53, 1.0],  $p = .021$ ,  $g = .20$ ), and  $60^\circ$  ( $\hat{p} = .67$  [.50, 1.0],  $p = .049$ ,  $g = .17$ ) levels of spatial separation, and marginally nonsignificant when the two harmonics were separated by  $15^\circ$  ( $\hat{p} = .63$  [.47, 1.0],  $p = .100$ ,  $g = .13$ ). A similar pattern of results was obtained when sample sizes were limited to include only those females ( $N = 25$ ) that were tested at all four levels of spatial separation ( $7.5^\circ$ :  $\hat{p} = .80$  [.62, 1.0],  $p = .002$ ,  $g = .30$ ;  $15^\circ$ :  $\hat{p} = .68$  [.496, 1.0],  $p = .054$ ,  $g = .18$ ;  $30^\circ$ :  $\hat{p} = .76$  [.58, 1.0],  $p = .007$ ,  $g = .26$ ; and  $60^\circ$ :  $\hat{p} = .76$  [.58, 1.0],  $p = .007$ ,  $g = .26$ ). The proportions of this smaller subset of females choosing the bimodal standard call over one of the harmonics in the spatially incoherent alternative did not differ across the four levels of spatial separation (Cochran's Q Test:  $Q = 1.08$ ,  $p = .783$ ,  $df = 3$ ). The observed proportions of subjects that chose the speakers broadcasting the spatially coherent standard call and the two spatially separated harmonic calls also did not differ from the null expected proportions (0.84, 0.0, or 0.16) based on the results of Experiment 1 (Fig. 2b–e; Table 2). Together, these results revealed two important findings. First, a spatially coherent bimodal call (i.e., the standard call) was preferred over a spatially incoherent alternative. And second, females chose the bimodal standard call and each of the two spatially separated single-harmonic stimuli in about the same relative proportions as expected based on the patterns of spectral preferences found in Experiment 1.

To what extent might the above-described pattern of preferences be attributable to a breakdown of source localizability with respect to the spatially incoherent alternative? A number of previous studies have shown that frogs exhibit highly accurate phonotaxis movements toward sources of real and synthetic advertisement calls (Gerhardt & Rheinlaender, 1980, 1982; Jorgensen & Gerhardt, 1991; Rheinlaender, Gerhardt, Yager, & Capranica, 1979; Shen et al., 2008; Ursprung, Ringler, & Hödl, 2009). Moreover, some treefrogs appear capable of discriminating between different angles of sound incidence at least as small as  $15^\circ$  (Klump & Gerhardt, 1989). Importantly, previous studies of treefrogs have also shown that female response times increase under conditions in which they have difficulty localizing the source of an otherwise attractive signal (Bee & Riemersma, 2008; Feng, et al., 1976). The accuracy of angular orientation toward conspecific calls can also be degraded in the presence of overlapping sounds (e.g., Marshall, Schwartz, & Gerhardt, 2006).

There was little evidence in the present study to suggest that response latency and angular orientation were affected by spatial incoherence. An examination of mean response latencies and 95% confidence intervals (Fig. 3) revealed significant overlap between those for choices of speakers broadcasting the spatially coherent calls of Test 1 in Experiment 1 and both spatially coherent and incoherent stimuli in the present experiment. One-way ANOVAs failed to reveal significant differences among  $\log_{10}$ -transformed mean latencies at all levels of angular separation:  $7.5^\circ$ ,  $F(2, 45) = .43$ ,  $p = .652$ ,  $\eta^2 = .02$ ;  $15^\circ$ ,  $F(2, 45) = .93$ ,  $p = .401$ ,  $\eta^2 = .04$ ;  $30^\circ$ ,  $F(2, 45) = .53$ ,  $p = .590$ ,  $\eta^2 = .02$ ; and  $60^\circ$ ,  $F(2, 46) = .13$ ,  $p = .882$ ,  $\eta^2 < .01$ . These small effects on response latency associated with females' choices of either spatially coherent or incoherent stimuli do not strongly support the idea that preferences for spatially coherent calls resulted from a breakdown of source localizability.

The mean vectors ( $\mu \pm 95\%$  CI) at which those subjects choosing one of the spatially incoherent, single-harmonic stimuli first touched the arena wall are depicted in Figure 4 relative to circular plots of the raw data. Across conditions, mean vectors ranged between  $-17^\circ$  and  $+22^\circ$  around the position of the chosen speaker (designated as  $0^\circ$ ). The corresponding median vectors for the  $7.5^\circ$ ,  $15^\circ$ ,  $30^\circ$ , and  $60^\circ$  conditions were  $0^\circ$ ,  $0^\circ$ ,  $15^\circ$ , and  $0^\circ$ , respectively. These results are similar to those obtained from subjects in Test 1 of Experiment 1, which chose between two spatially coherent bimodal calls. In that test, the

mean vector angle relative to the chosen speaker ( $0^\circ$ ) was  $-4^\circ$  [ $-24^\circ$ ,  $16^\circ$ ]; the median vector was  $0^\circ$ . As illustrated in Figure 4, the distributions of angles at which subjects first touched the wall relative to a chosen, spatially incoherent stimulus in Experiment 2 and the spatially coherent standard call in Experiment 1 did not differ significantly at any level of angular separation:  $7.5^\circ$ ,  $\chi^2(8) = 6.0$ ,  $p = .648$ ;  $15^\circ$ ,  $\chi^2(10) = 9.8$ ,  $p = .457$ ;  $30^\circ$ ,  $\chi^2(9) = 15.4$ ,  $p = .081$ ; and  $60^\circ$ ,  $\chi^2(9) = 9.4$ ,  $p = .404$ . Consistent with the analyses of response latency, these circular statistical analyses indicate that interpretations of the results in terms of a breakdown of source localizability are not strongly supported by the currently available data.

Taken together, the results of Experiment 2 suggest that females of *H. chrysoscelis* perceptually and behaviorally discriminated between stimuli in which the two harmonics composing a bimodal call were either spatially coherent or spatially incoherent. They preferred spatially coherent calls. In addition, the level of discrimination was similar across all four levels of  $\theta$ , ranging from  $60^\circ$  spatial separation down to just  $7.5^\circ$ . These data are broadly consistent with the hypothesis that spatial coherence could function as a cue for simultaneous integration in frogs. According to this interpretation of the data, the preferences exhibited for the spatially coherent standard call in Experiment 2 represent a breakdown of simultaneous integration that occurred when the two harmonics originated from different locations.

It is important to consider what acoustic cues might have been available for females in discriminating between spatially coherent and spatially incoherent calls. Clearly one set of available cues is the typical binaural cues used in sound localization by frogs, which include the differences in level and arrival time between sounds originating from different locations (reviewed in Feng & Schellart, 1999; Gerhardt & Bee, 2007; Gerhardt & Huber, 2002). In a strict sense, these differences represent the cues usually considered in discussions of common spatial origin as an auditory grouping cue (Darwin, 1997, 2006; Darwin & Carlyon, 1995). At present, however, we cannot conclude with certainty that the subjects in this study were using these specific cues (in part or exclusively). This follows because the physical relationships between subjects and the two spatially separated speakers were not fixed during a test. Rather, subjects were *required* to move about in the arena to demonstrate discriminative behavioral responses. This methodological requirement, in turn, made available at least three additional cues that might have contributed to the observed patterns of discrimination between spatially coherent and incoherent calls.

Two of these additional cues are the onset times and ongoing phases of the two spatially separated harmonics. At the subject release site, the two harmonics had a common onset and phase (see Supplemental Material). Disparities in these two cues, however, would have resulted from differences in the path lengths travelled by the two sounds as subjects moved about in the test arena. In the present study, the maximum disparity in onset times would have occurred when a subject was sitting directly in front of one of the two spatially separated speakers. In this situation, the magnitude of difference in the arrival times of the two harmonics would vary from a maximum of about 3 ms in the  $60^\circ$  condition to a minimum of about 0.4 ms in the  $7.5^\circ$  condition (based on using 343.2 m/s as the speed of sound in air at  $20^\circ\text{C}$  and ignoring the rather small external binaural disparities). Common onsets are known to be a potent cue in simultaneous integration by humans (Darwin, 1997, 2006; Darwin & Carlyon, 1995), but their saliency as an auditory grouping cue in frogs has never been directly tested.

Variation in the path lengths of sound propagation can also introduce variation in the relative phases of harmonically-related sounds. For example, the magnitude of arrival time disparities described above translate into phase differences between the two harmonics in an ongoing pulse of between  $68^\circ$  and  $159^\circ$  in the  $60^\circ$  and  $7.5^\circ$  conditions, respectively (based

on using a period of 0.91 ms for the 1.1 kHz fundamental frequency). While some frogs may be capable of discriminating between harmonic stimuli differing in phase spectrum (e.g., Hainfeld, Boatright-Horowitz, Boatright-Horowitz, & Simmons, 1996), it is presently unknown whether such phase differences would interfere with selective phonotaxis. It would be somewhat surprising, however, if variation in the relative phases of the two harmonics experienced during phonotaxis had large influences on determining the attractiveness of a male's advertisement call. In the reverberant and structurally complex habitats in which these frogs breed, natural and complex variation in the relative phases of the two harmonics is almost certainly introduced at the positions of a nonstationary receiver as a result of variation in frequency-dependent scattering and reflections (e.g., by vegetation). Thus, it seems unlikely that introduced phase differences could explain the pattern of female preferences observed in this study.

A third cue stemming from a subject's freedom to move within the sound field involves changes in the relative amplitudes of the two spatially separated harmonics. Consider, for example, a subject that moved away from the central release point toward a speaker broadcasting one of the harmonics in the spatially incoherent bimodal call. During the approach, the amplitude of the harmonic coming from the approached speaker would increase along a steeper gradient relative to that of the other spatially separated harmonic. This effect of different sound pressure gradients would result in variation in the relative amplitudes of the two simultaneous but spatially separated harmonics. The extent to which this introduced variation might contribute to an ability to discriminate between spatially coherent and incoherent stimuli remains unknown. However, as with the relative phases of the two harmonics, we might also expect females to be somewhat permissive of variation in the relative amplitudes of the two harmonics. In the animal's natural habitat, the two harmonics are likely degraded and attenuated to different degrees as a function of distance, for example, due to wavelength-dependent effects of microhabitat and the excess attenuation of higher frequencies during sound propagation (Bradbury & Vehrencamp, 1998; Forrest, 1994). Thus, in the natural listening environment of a chorus, females almost certainly experience some variation in the relative amplitudes of the harmonics composing the calls of males they approach and eventually select as mates. It seems unlikely that such variation would have strong influences on the attractiveness of a signal to a female. The empirical results of two-choice phonotaxis tests manipulating the relative amplitudes of harmonics in alternative stimuli generally support this assertion (H. C. Gerhardt, personal communication, November 27, 2009).

It will be important in future studies to empirically demonstrate which specific cue or cues are the most salient in allowing frogs to discriminate between spatially coherent and incoherent stimuli. Based on results of this study, it is impossible to distinguish between the contribution of binaural cues and cues related to differences in the onset times, relative phases, and relative amplitudes of the two harmonics. However, considerations of acoustic signal perception in structurally complex habitats suggest that cues related to binaural differences or onset times are probably more likely to be involved in promoting simultaneous integration than differences in relative phase and amplitude.

## General Discussion

There is considerable interest in understanding how nonhuman animals solve the everyday problems of auditory scene analysis that they encounter in their natural environments (reviews in Bee & Micheyl, 2008; Feng & Ratnam, 2000; Hulse, 2002). Most often, this interest centers around the perception of acoustic signals (or their echoes). A number of relatively recent studies have shown that auditory scene analysis in some form operates in the perception of acoustic signals by animals as diverse as katydids (Schul & Sheridan,

2006), frogs (Simmons & Bean, 2000), penguins (Aubin & Jouventin, 1998), songbirds (reviewed in Hulse, 2002), monkeys (Miller, Dibble, & Hauser, 2001; Petkov, O'Connor, & Sutter, 2003; Weiss & Hauser, 2002), mice (Geissler & Ehret, 2002), and bats (Barber, Razak, & Fuzessery, 2003; Kanwal, Medvedev, & Micheyl, 2003; Moss & Surlykke, 2001). These studies and others notwithstanding, the general importance of the various potential cues for auditory grouping still remain largely unexplored in studies of animal acoustic communication (Bee & Micheyl, 2008).

In humans, harmonicity (i.e., common F0) and common onsets are two of the most potent acoustic cues that promote the grouping of simultaneous sounds across frequency (reviewed in Darwin, 1997, 2006; Darwin & Carlyon, 1995). There are a few studies to suggest that these two auditory grouping cues may also operate in acoustic communication by nonhuman animals. For example, harmonicity influences the perception of harmonically rich vocalizations in cotton-top tamarins, *Saguinus oedipus* (Weiss & Hauser, 2002), North American bullfrogs, *Rana catesbeiana* (Simmons & Bean, 2000), and green treefrogs, *Hyla cinerea* (Simmons, 1988; but see Gerhardt, Allan, & Schwartz, 1990). Geissler and Ehret (2002) identified the common onset of simultaneous harmonics as a grouping cue important in the recognition of pup wriggling calls by nursing mice (*Mus domesticus*). Results from the present study are unable to rule out the hypothesis that cues related to spatial coherence promote integration of the simultaneous spectral components that compose the advertisement call and that are primarily encoded by physically distinct sensory papillae. Somewhat surprisingly, departures from common spatial origin ranging between 7.5° and 60° angular separation yielded equivalent patterns of female preferences for a spatially coherent bimodal call. These results offer some interesting contrasts with what might have been predicted based on previous studies of simultaneous (across frequency) integration in humans and sequential (across time) integration in frogs.

Despite our often strong intuition that grouping sounds originating from the same location would be a good strategy for assigning them to the same source, empirical evidence suggests that common spatial origin is a relatively weak cue for simultaneous integration in human hearing and speech perception (see reviews in Darwin, 1997, 2006; Darwin & Carlyon, 1995). For example, a number of human studies have now shown that cues related to common spatial origin (e.g., inter-aural time differences) have relatively weak effects on simultaneous integration compared to other cues, such as harmonicity or onset synchrony. Moreover, other (nonspatial) cues for simultaneous grouping can actually override grouping by common location and even promote the perceptual integration of simultaneous sounds presented to different ears. Current evidence from humans is thus consistent with the idea that the across-frequency grouping of simultaneous sounds, for example based on their harmonicity and common onsets, occurs prior to their joint assignment to a particular location (Darwin, 1997, 2006; Darwin & Carlyon, 1995).

The finding that an angular separation as small as 7.5° appeared to interfere with simultaneous integration in gray treefrogs represents a stark contrast with previous investigations into the role of spatial coherence in sequential integration by frogs. Current evidence suggests that frogs are rather permissive of departures from spatial coherence in the sequential integration of sounds across time (instead of across frequency). In recent studies of the túngara frog (*Physalaemus pustulosus*), Farris, Rand, and Ryan (2002, 2005) showed that females perceptually group the two sequential components composing complex calls (the 'whine' and the 'chuck') over angular separations of up to 135°. Likewise, Schwartz and Gerhardt (1995) and Bee and Riemersma (2008), studying *H. versicolor* and *H. chrysoscelis*, respectively, have shown that female gray treefrogs continue to group sequential sounds (e.g., interleaved pulse sequences) when they are broadcast at angular separations ranging between 45° and 180°. Males of the Australian quacking frog (*Crinia*

*georgiana*) group the temporally separate notes composing multiple-note calls at angular separations of up to 180° (Gerhardt, Roberts, Bee, & Schwartz, 2000).

Aside from constraints posed by the physics of sound, there are relatively few *a priori* reasons to expect that frogs and humans *must* solve similar tasks of simultaneous or sequential integration in similar ways. After all, frogs are unique among vertebrates in having two physically distinct sensory papillae that encode different frequency ranges of airborne sound. Unlike humans and other mammals, frogs use a pressure-gradient system for sound localization, similar to some birds (Larsen, Dooling, & Michelsen, 2006) and insects (Gerhardt & Huber, 2002). In frogs, this system involves a coupling of the tympanic middle ears on opposite sides of the head through permanently open Eustachian tubes. Moreover, there are extra-tympanic pathways by which sound energy may be transmitted to the inner ears of frogs, such as through the body wall and lungs (reviewed in Gerhardt & Bee, 2007; Mason, 2007). How these features of the frog auditory system might either facilitate or constrain the use of various potential cues for auditory grouping, including common spatial origin, represent questions for future study, as they have not been explored in any depth. One important next step in understanding sequential and simultaneous integration in frogs will be to assess the salience and relative importance of different cues by systematically placing these cues in conflict with each other.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## References

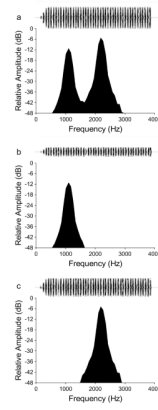
- Aubin T, Jouventin P. Cocktail-party effect in king penguin colonies. *Proceedings of the Royal Society of London Series B-Biological Sciences* 1998;265(1406):1665–1673.
- Barber JR, Razak KA, Fuzessery ZM. Can two streams of auditory information be processed simultaneously? Evidence from the gleaning bat *Antrozous pallidus*? *Journal of Comparative Physiology A* 2003;189(11):843–855.
- Batschelet, E. *Circular Statistics in Biology*. London: Academic Press; 1981.
- Bee MA. Sound source segregation in grey treefrogs: Spatial release from masking by the sound of a chorus. *Animal Behaviour* 2007;74:549–558.
- Bee MA, Micheyl C. The cocktail party problem: What is it? How can it be solved? And why should animal behaviorists study it? *Journal of Comparative Psychology* 2008;122:235–251. [PubMed: 18729652]
- Bee MA, Riemersma KK. Does common spatial origin promote the auditory grouping of temporally separated signal elements in grey treefrogs? *Animal Behaviour* 2008;76:831–843. [PubMed: 19727419]
- Bee MA, Swanson EM. Auditory masking of anuran advertisement calls by road traffic noise. *Animal Behaviour* 2007;74:1765–1776.
- Bradbury, JW.; Vehrencamp, SL. *Principles of Animal Communication*. Sunderland, MA: Sinauer Associates; 1998.
- Bregman, AS. *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, MA: MIT Press; 1990.
- Brumm H, Slabbekoorn H. Acoustic communication in noise. *Advances in the Study of Behavior* 2005;35:151–209.
- Capranica, RR.; Moffat, JM. Neurobehavioral correlates of sound communication in anurans. In: Ewert, JP.; Capranica, RR.; Ingle, DJ., editors. *Advances in Vertebrate Neuroethology*. New York: Plenum Press; 1983. p. 701-730.
- Carlyon RP. How the brain separates sounds. *Trends in Cognitive Sciences* 2004;8(10):465–471. [PubMed: 15450511]



- Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*. 2. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
- Darwin CJ. Auditory grouping. *Trends in Cognitive Sciences* 1997;1(9):327–333.
- Darwin CJ. Contributions of binaural information to the separation of different sound sources. *International Journal of Audiology* 2006;45:S20–S24. [PubMed: 16938771]
- Darwin, CJ.; Carlyon, RP. Auditory grouping. In: Moore, BCJ., editor. *Hearing*. New York: Academic Press; 1995. p. 387–424.
- Erdfelder E, Faul F, Buchner A. GPOWER: A general power analysis program. *Behavior Research Methods Instruments & Computers* 1996;28(1):1–11.
- Farris HE, Rand AS, Ryan MJ. The effects of spatially separated call components on phonotaxis in túngara frogs: Evidence for auditory grouping. *Brain Behavior and Evolution* 2002;60(3):181–188.
- Farris HE, Rand AS, Ryan MJ. The effects of time, space and spectrum on auditory grouping in túngara frogs. *Journal of Comparative Physiology A* 2005;191(12):1173–1183.
- Faul F, Erdfelder E, Lang AG, Buchner A. G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods* 2007;39(2):175–191. [PubMed: 17695343]
- Fay RR, Popper AN. Evolution of hearing in vertebrates: The inner ears and processing. *Hearing Research* 2000;149(1–2):1–10. [PubMed: 11033242]
- Feng AS, Gerhardt HC, Capranica RR. Sound localization behavior of the green treefrog (*Hyla cinerea*) and the barking treefrog (*Hyla gratiosa*). *Journal of Comparative Physiology* 1976;107(3):241–252.
- Feng AS, Ratnam R. Neural basis of hearing in real-world situations. *Annual Review of Psychology* 2000;51:699–725.
- Feng, AS.; Schellart, NAM. Central auditory processing in fish and amphibians. In: Fay, RR.; Popper, AN., editors. *Comparative Hearing: Fish and Amphibians*. New York: Springer; 1999. p. 218–268.
- Feng, AS.; Schul, J. Sound processing in real-world environments. In: Narins, PA.; Feng, AS.; Fay, RR.; Popper, AN., editors. *Hearing and Sound Communication in Amphibians*. New York: Springer; 2007. p. 323–350.
- Forrest TG. From sender to receiver: Propagation and environmental effects on acoustic signals. *American Zoologist* 1994;34(6):644–654.
- Fuzessery ZM, Feng AS. Frequency selectivity in the anuran auditory midbrain: Single unit responses to single and multiple tone stimulation. *Journal of Comparative Physiology* 1982;146(4):471–484.
- Geissler DB, Ehret G. Time-critical integration of formants for perception of communication calls in mice. *Proceedings of the National Academy of Sciences of the United States of America* 2002;99(13):9021–9025. [PubMed: 12070345]
- Gerhardt HC. Sound pressure levels and radiation patterns of vocalizations of some North American frogs and toads. *Journal of Comparative Physiology* 1975;102(1):1–12.
- Gerhardt HC. 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* 1981;144(1):9–16.
- Gerhardt, HC. Phonotaxis in female frogs and toads: execution and design of experiments. In: Klump, GM.; Dooling, RJ.; Fay, RR.; Stebbins, WC., editors. *Methods in Comparative Psychoacoustics*. Basel: Birkhäuser Verlag; 1995. p. 209–220.
- Gerhardt HC. Acoustic communication in two groups of closely related treefrogs. *Advances in the Study of Behavior* 2001;30:99–167.
- Gerhardt HC. Acoustic spectral preferences in two cryptic species of grey treefrogs: Implications for mate choice and sensory mechanisms. *Animal Behaviour* 2005;70:39–48.
- Gerhardt HC, Allan S, Schwartz JJ. Female green treefrogs (*Hyla cinerea*) do not selectively respond to signals with a harmonic structure in noise. *Journal of Comparative Physiology A* 1990;166(6):791–794.
- Gerhardt, HC.; Bee, MA. Recognition and localization of acoustic signals. In: Narins, PM.; Feng, AS.; Fay, RR.; Popper, AN., editors. *Hearing and Sound Communication in Amphibians*. Vol. 28. New York: Springer; 2007. p. 113–146.

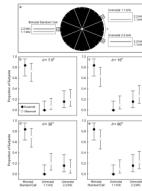
- Gerhardt HC, Doherty JA. Acoustic communication in the gray treefrog, *Hyla versicolor*: Evolutionary and neurobiological implications. *Journal of Comparative Physiology A* 1988;162(2):261–278.
- Gerhardt, HC.; Huber, F. *Acoustic Communication in Insects and Anurans: Common Problems and Diverse Solutions*. Chicago: Chicago University Press; 2002.
- Gerhardt HC, Martinez-Rivera CC, Schwartz JJ, Marshall VT, Murphy CG. Preferences based on spectral differences in acoustic signals in four species of treefrogs (Anura : Hylidae). *Journal of Experimental Biology* 2007;210(17):2990–2998. [PubMed: 17704074]
- Gerhardt HC, Rheinlaender J. Accuracy of sound localization in a miniature dendrobatid frog. *Naturwissenschaften* 1980;67(7):362–363.
- Gerhardt HC, Rheinlaender J. Localization of an elevated sound source by the green tree frog. *Science* 1982;217(4560):663–664. [PubMed: 17817539]
- Gerhardt HC, Roberts JD, Bee MA, Schwartz JJ. Call matching in the quacking frog (*Crinia georgiana*). *Behavioral Ecology and Sociobiology* 2000;48(3):243–251.
- Gerhardt, HC.; Schwartz, JJ. Auditory tuning, frequency preferences and mate choice in anurans. In: Ryan, MJ., editor. *Anuran Communication*. Washington DC: Smithsonian Institution Press; 2001. p. 73-85.
- Gerhardt HC, Tanner SD, Corrigan CM, Walton HC. Female preference functions based on call duration in the gray tree frog (*Hyla versicolor*). *Behavioral Ecology* 2000;11(6):663–669.
- Hainfeld CA, Boatright-Horowitz SL, Boatright-Horowitz SS, Simmons AM. Discrimination of phase spectra in complex sounds by the bullfrog (*Rana catesbeiana*). *Journal of Comparative Physiology a-Sensory Neural and Behavioral Physiology* 1996;179(1):75–87.
- Holloway AK, Cannatella DC, Gerhardt HC, Hillis DM. Polyploids with different origins and ancestors form a single sexual polyploid species. *American Naturalist* 2006;167(4):E88–E101.
- Hulse SH. Auditory scene analysis in animal communication. *Advances in the Study of Behavior* 2002;31:163–200.
- Jorgensen MB, Gerhardt HC. Directional hearing in the gray tree frog *Hyla versicolor*: Eardrum vibrations and phonotaxis. *Journal of Comparative Physiology A* 1991;169(2):177–183.
- Kanwal JS, Medvedev AV, Micheyl C. Neurodynamics for auditory stream segregation: Tracking sounds in the mustached bat's natural environment. *Network: Computation in Neural Systems* 2003;14(3):413–435.
- Klump GM, Gerhardt HC. Sound localization in the barking treefrog. *Naturwissenschaften* 1989;76(1):35–37. [PubMed: 2704420]
- Langemann, U.; Klump, GM. Perception and acoustic communication networks. In: McGregor, PK., editor. *Animal Communication Networks*. Cambridge: Cambridge University Press; 2005. p. 451-480.
- Larsen ON, Dooling RJ, Michelsen A. The role of pressure difference reception in the directional hearing of budgerigars (*Melopsittacus undulatus*). *Journal of Comparative Physiology A* 2006;192(10):1063–1072.
- Lewis, ER.; Narins, PM. The acoustic periphery of amphibians: Anatomy and physiology. In: Fay, RR.; Popper, AN., editors. *Hearing: Fish and Amphibians*. Vol. 11. New York: Springer; 1999. p. 101-154.
- Marshall VT, Schwartz JJ, Gerhardt HC. Effects of heterospecific call overlap on the phonotactic behaviour of grey treefrogs. *Animal Behaviour* 2006;72:449–459.
- Mason, MJ. Pathways for sound transmission to the inner ear in amphibians. In: Narins, PM.; Feng, AS.; Fay, RR.; Popper, AN., editors. *Hearing and Sound Communication in Amphibians*. Vol. 28. New York: Springer; 2007. p. 147-183.
- Miller CT, Dibble E, Hauser MD. Amodal completion of acoustic signals by a nonhuman primate. *Nature Neuroscience* 2001;4(8):783–784.
- Moss CF, Surlykke A. Auditory scene analysis by echolocation in bats. *Journal of the Acoustical Society of America* 2001;110(4):2207–2226. [PubMed: 11681397]
- Narins, PM.; Zelick, R. The effects of noise on auditory processing and behavior in amphibians. In: Fritsch, B.; Ryan, MJ.; Wilczynski, W.; Hetherington, TE.; Walkowiak, W., editors. *The Evolution of the Amphibian Auditory System*. New York: Wiley & Sons; 1988. p. 511-536.

- Petkov CI, O'Connor KN, Sutter ML. Illusory sound perception in macaque monkeys. *Journal of Neuroscience* 2003;23(27):9155–9161. [PubMed: 14534249]
- Ptacek MB, Gerhardt HC, Sage RD. Speciation by polyploidy in treefrogs: Multiple origins of the tetraploid, *Hyla versicolor*. *Evolution* 1994;48(3):898–908.
- Rheinlaender J, Gerhardt HC, Yager DD, Capranica RR. Accuracy of phonotaxis by the green treefrog (*Hyla cinerea*). *Journal of Comparative Physiology* 1979;133(4):247–255.
- Schul J, Sheridan RA. Auditory stream segregation in an insect. *Neuroscience* 2006;138(1):1–4. [PubMed: 16378693]
- Schwartz JJ, Gerhardt HC. Directionality of the auditory system and call pattern recognition during acoustic interference in the gray treefrog, *Hyla versicolor*. *Auditory Neuroscience* 1995;1:195–206.
- Shen JX, Feng AS, Xu ZM, Yu ZL, Arch VS, Yu XJ, et al. Ultrasonic frogs show hyperacute phonotaxis to female courtship calls. *Nature* 2008;453(7197):914–U946. [PubMed: 18469804]
- Simmons AM. Selectivity for harmonic structure in complex sounds by the green treefrog (*Hyla cinerea*). *Journal of Comparative Physiology A* 1988;162(3):397–403.
- Simmons AM, Bean ME. Perception of mistuned harmonics in complex sounds by the bullfrog (*Rana catesbeiana*). *Journal of Comparative Psychology* 2000;114(2):167–173. [PubMed: 10890588]
- StatSoft. STATISTICA (data analysis software system), version 7.1. 2006. www.statsoft.com
- Ursprung E, Ringler M, Hödl W. Phonotactic approach pattern in the neotropical frog *Allobates femoralis*: A spatial and temporal analysis. *Behaviour* 2009;146:153–170.10.1163/156853909x410711
- Weiss DJ, Hauser MD. Perception of harmonics in the combination long call of cottontop tamarins, *Saguinus oedipus*. *Animal Behaviour* 2002;64:415–426.
- Wells, KD.; Schwartz, JJ. The behavioral ecology of anuran communication. In: Narins, PM.; Feng, AS.; Fay, RR.; Popper, AN., editors. *Hearing and Sound Communication in Amphibians*. Vol. 28. New York: Springer; 2007. p. 44-86.
- Zakon, HH.; Wilczynski, W. The physiology of the anuran eighth nerve. In: Fritsch, B.; Wolkowiak, W.; Ryan, MJ.; Wilczynski, W.; Hetherington, T., editors. *The Evolution of the Amphibian Auditory System*. New York: Wiley; 1988. p. 125-155.
- Zar, JH. *Biostatistical Analysis*. 4. Upper Saddle River, NJ: Prentice-Hall; 1999.



**Figure 1.**

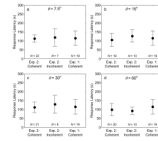
Acoustic stimuli in Experiments 1 and 2. Power spectra (512 point FFT, Hanning windows) showing (a) the synthetic bimodal standard call and unimodal calls consisting of either (b) the 1.1 kHz harmonic alone or (c) the 2.2 kHz harmonic alone at the same relative amplitudes as in the standard call depicted in (a). The waveform of each stimulus is depicted above its power spectrum to illustrate further the differences in relative amplitude between the two harmonics.



**Figure 2.**

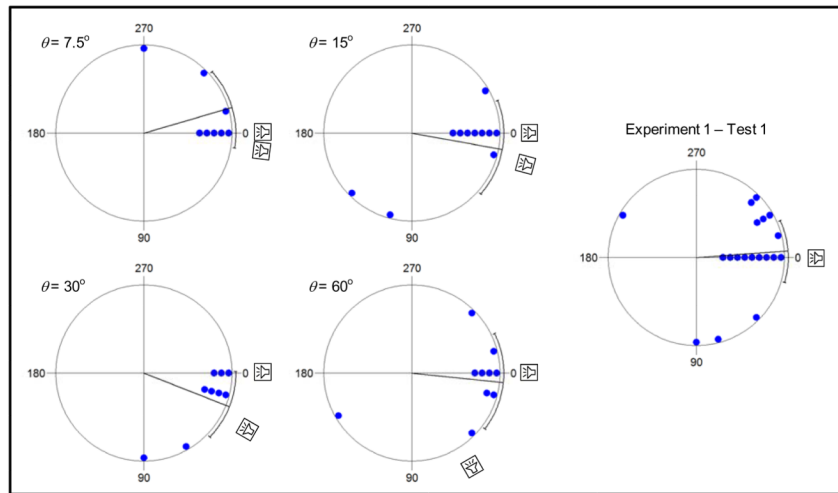
Design and results of Experiment 2. (a) Schematic diagram illustrating the positions of the spatially coherent, bimodal standard call and the two spatially separated unimodal stimuli composing the spatially incoherent bimodal alternative. The angle of separation ( $\theta$ ) between the two single-harmonic stimuli was  $7.5^\circ$ ,  $15^\circ$ ,  $30^\circ$ , or  $60^\circ$  in separate tests. (b–e) Results from phonotaxis experiments showing the observed (open circles) proportions of females that approached the bimodal standard call or one of the single-harmonic alternatives. Expected proportions based on Experiment 1 are depicted as black circles. Error bars depict exact 95% binomial proportion confidence intervals (Zar, 1999).





**Figure 3.**

Latencies in response to the spatially coherent and incoherent bimodal calls. Points and error bars depict the mean and 95% confidence intervals. In each plot, the left and center points depict response latencies for subjects in Experiment 2 that chose either the bimodal standard call (Exp. 2: Coherent) or one of the two spatially separated speakers broadcasting the spatially incoherent bimodal alternative (Exp.2: Incoherent); in each plot, the right-most point depicts response latencies for subjects from Test 1 of Experiment 1, which were given a choice between two spatially coherent bimodal calls (Exp. 1: Coherent).



**Figure 4.** Angular orientation in response to the spatially incoherent stimuli of Experiment 2. For four levels of angular separation ( $7.5^\circ$ ,  $15^\circ$ ,  $30^\circ$ , and  $60^\circ$ ), each point depicts the angle at which those subjects that chose a speaker broadcasting one of the harmonics in the spatially incoherent bimodal call first touched the wall of the circular test arena. The positions of points are depicted relative to the position of the speaker that was eventually chosen (designated  $0^\circ$ ) with the direction of the other spatially separated speaker toward  $90^\circ$ . For comparison, similar data are also shown for the responses of subjects to the spatially coherent standard call in Test 1 of Experiment 1. Also shown are the mean vectors (lines) and their 95% confidence intervals (brackets).

Table 1

Numbers (N) and Proportions ( $\hat{p}$ ) of Females Choosing Each Alternative in Experiment 1.

Test	Alternative 1	Alternative 2	Alternative 1		Alternative 2		Binomial $p$	Effect size ( $g$ )
			N	$\hat{p}$ [95% CI]	N	$\hat{p}$ [95% CI]		
1	Standard Call	Standard Call	11	.55 [.32, .77]	9	.45 [.23, .68]	.824	.05
2	Standard Call	1.1 kHz (Rel. Amp.)	20	1.00 [.83, 1.0]	0	.00 [0, .17]	< .001	.50
3 <sup>a</sup>	Standard Call	2.2 kHz (Rel. Amp.)	21	.84 [.63, .95]	4	.16 [.05, .36]	< .001	.34
4	Standard Call	1.1 kHz (Full Amp.)	20	1.0 [.83, 1.0]	0	.00 [0, .17]	< .001	.50
5 <sup>a</sup>	Standard Call	2.2 kHz (Full Amp.)	20	.67 [.47, .83]	10	.33 [.17, .53]	.099	.17
6	1.1 kHz (Full Amp.)	2.2 kHz (Full Amp.)	0	.00 [0, .17]	20	1.00 [.83, 1.0]	< .001	.50

Note. "Rel. Amp." and "Full Amp." designate the amplitude of the unimodal alternative as either the same as that in the standard call ("Rel. Amp.") or 75 dB SPL ("Full Amp.").

<sup>a</sup>The proportions of subjects choosing Alternatives 1 and 2 at a sample size of  $N = 20$  (i.e., the first 20 females to make a choice) were as follows. Test 3: Alternative 1 ( $\hat{p} = .85$  [.62, .97]) versus Alternative 2 ( $\hat{p} = .15$  [.03, .38]); Test 5: Alternative 1 ( $\hat{p} = .70$  [.45, .88]) versus Alternative 2 ( $\hat{p} = .30$  [.12, .54]).

**Table 2**

Results of Z Tests Comparing the Expected and Observed Proportions Choosing the Spatially Coherent Standard Call and the Two Harmonics in the Spatially Incoherent Bimodal Call in Experiment 2.

Angular Separation	Choice	<i>z</i>	<i>p</i>	<i>h</i>
7.5°	Standard Call	.66	.509	.26
	1.1 kHz	.50	.617	.52
	2.2 kHz	.03	.976	.10
15°	Standard Call	1.44	.150	.48
	1.1 kHz	.85	.395	.64
	2.2 kHz	.66	.509	.26
30°	Standard Call	.90	.368	.34
	1.1 kHz	1.69	.091	.93
	Standard Call	.26	.795	.18
60°	Standard Call	1.13	.258	.41
	1.1 kHz	.85	.395	.64
	2.2 kHz	.31	.757	.19

Note. Expected proportions are derived from the outcome of Test 1 in Experiment 1 (reported in Table 1). The corresponding raw data are depicted in Figure 2.