Pulse-number discrimination by Cope's gray treefrog (*Hyla chrysoscelis*) in modulated and unmodulated noise

Alejandro Vélez^{a)} and Betsy Jo Linehan-Skillings

Department of Ecology, Evolution and Behavior, University of Minnesota, St. Paul, Minnesota 55108

Yuwen Gu and Yuting Sun

School of Statistics, University of Minnesota, Minneapolis, Minnesota 55455

Mark A. Bee

Department of Ecology, Evolution and Behavior, University of Minnesota, St. Paul, Minnesota 55108

(Received 22 March 2013; revised 25 July 2013; accepted 9 August 2013)

In Cope's gray treefrog (*Hyla chrysoscelis*), thresholds for recognizing conspecific calls are lower in temporally modulated noise backgrounds compared with unmodulated noise. The effect of modulated noise on discrimination among different conspecific calls is unknown. In quiet, females prefer calls with relatively more pulses. This study tested the hypotheses that noise impairs selectivity for longer calls and that processes akin to dip listening in modulated noise can ameliorate this impairment. In two-stimulus choice tests, female subjects were allowed to choose between an average-length call and a shorter or longer alternative. Tests were replicated at two signal levels in quiet and in the presence of chorus-shaped noise that was unmodulated, modulated by a sinusoid, or modulated by envelopes resembling natural choruses. When subjects showed a preference, it was always for the relatively longer call. Noise reduced preferences for longer calls, but the magnitude of this reduction was unrelated to whether the noise envelope was modulated or unmodulated. Together, the results are inconsistent with the hypothesis that dip listening improves a female gray treefrog's ability to select longer calls in modulated compared with unmodulated noise.

© 2013 Acoustical Society of America. [http://dx.doi.org/10.1121/1.4820883]

PACS number(s): 43.80.Nd, 43.80.Lb, 43.66.Gf [MJO]

Pages: 3079–3089

I. INTRODUCTION

The noise and acoustic clutter generated in breeding choruses of frogs present receivers with a communication problem analogous to the human “cocktail party problem” (Cherry, 1953; Bee and Micheyl, 2008; Bee, 2012; Vélez *et al.*, 2014). The vocalizations of many frogs reach or exceed peak sound pressure levels of 90–100 dB SPL (re 20 μ Pa at 1 m) (Gerhardt, 1975). The sustained noise levels in large, dense choruses of calling frogs can be quite high (Narins, 1982; Schwartz *et al.*, 2001) and choruses can be audible to humans over large distances (e.g., 2 km; Arak, 1983). Chorus noise and overlapping calls can decrease a signal's active space, impair species recognition, interfere with call-type discrimination, and constrain the expression of female preferences for males producing certain types of calls (reviewed in Bee, 2012; Vélez *et al.*, 2014). These challenges notwithstanding, female frogs nevertheless select appropriate mates based on analyses of the spectral and temporal properties of acoustic signals under the noisy conditions in a chorus (reviewed in Gerhardt and Huber, 2002).

There is increasing evidence to suggest frogs and humans rely on many of the same perceptual features of complex acoustic scenes in solving cocktail-party-like communication problems (reviewed in Bee, 2012; Vélez *et al.*,

2014). One such feature is temporal fluctuation in the level of background noise. Studies of “dip listening” (Buus, 1985) in humans show that speech-identification performance is better, and speech-recognition thresholds are lower, in the presence of temporally fluctuating speech and speech-shaped noise than in the presence of unmodulated noise (Festen and Plomp, 1990; Gustafsson and Arlinger, 1994; Bacon *et al.*, 1998; Füllgrabe *et al.*, 2006). These effects are attributed to the auditory system's ability to catch brief acoustic glimpses of speech when the noise momentarily “dips” to a low level. Dip listening constitutes one way by which the human auditory system ameliorates the problem of communicating in noisy environments, though it should be noted that hearing-impaired listeners experience relatively less masking release in fluctuating noise (e.g., Jin and Nelson, 2006). Questions about the extent to which dip listening enables nonhuman animals to solve analogous communication problems have so far received little attention. The present study investigated the effects of modulated and unmodulated noise on the ability of females of Cope's gray treefrog (*Hyla chrysoscelis*) to discriminate among conspecific vocalizations differing along a biologically relevant continuum.

A. Pulse-number discrimination

Males of *H. chrysoscelis* produce pulsatile advertisement calls (Gerhardt, 2001). The rate of pulses in the call is a critical cue for species recognition (Schul and Bush, 2002). In the context of sexual selection, the number of pulses in a call varies reliably among males and is an important acoustic

^{a)} Author to whom correspondence should be addressed. Current address: Department of Biological Sciences, Purdue University, 915 West State Street, West Lafayette, IN 47907. Electronic mail: avelezme@purdue.edu

cue females use to discriminate among potential mates (Gerhardt, 1994; Gerhardt *et al.*, 1996; Gerhardt *et al.*, 2000; Schwartz *et al.*, 2001; Bee, 2008; Ward *et al.*, 2013). Pulse number functions as a potentially honest indicator of a male's genetic quality. Males of a closely related sister species (the eastern gray treefrog, *H. versicolor*) that produce calls with more pulses also sire higher-quality offspring (Welch *et al.*, 1998). Hence, in evolutionary terms, there is a potential fitness advantage to females who mate with males producing longer calls.

In two-stimulus choice tests conducted in quiet, females of both Cope's and eastern gray treefrogs discriminate between two alternative calls differing only in pulse number, preferentially choosing longer calls with relatively more pulses (Gerhardt, 1994; Gerhardt *et al.*, 1996; Gerhardt *et al.*, 2000; Schwartz *et al.*, 2001; Bee, 2008; Ward *et al.*, 2013). Both species also exhibit a nonlinear preference for higher pulse numbers, in which discrimination against shorter calls is stronger when both alternatives are relatively short compared to when both alternatives are of average length or longer. Although this nonlinearity would at first appear consistent with Weber's Law, preferences depend on both the relative and absolute pulse numbers in alternative calls (Gerhardt *et al.*, 2000). Most relevant to the current study is work showing that "chorus-shaped noise" (i.e., noise with the frequency spectrum of natural choruses) reduces some female preferences for longer calls (Bee, 2008; Ward *et al.*, 2013). Thus, chorus noise potentially functions to limit the expression of adaptive female preferences for longer calls (e.g., Schwartz *et al.*, 2001).

B. Dip listening in modulated noise

The sound generated by choruses of Cope's gray treefrogs simultaneously fluctuates at slow rates ($< \sim 2$ Hz) and at a faster rate (~ 40 – 50 Hz). The slow rate of fluctuation appears to correspond to the call rates and call timing interactions of males in the chorus. The faster rate corresponds to the pulse rate of the advertisement call (Vélez and Bee, 2010). Evidence from studies that have measured masked signal-recognition thresholds (Bee and Schwartz, 2009) indicates that thresholds for responding to conspecific calls are significantly lower in temporally modulated, chorus-shaped noise than in unmodulated noise (Vélez and Bee, 2011, 2013; Schwartz *et al.*, 2013). Thus, gray treefrogs benefit from dip listening in temporally fluctuating noise in terms of improved signal recognition.

The present study extends this earlier work by investigating whether dip listening in modulated noise might also ameliorate the negative impacts of chorus noise on pulse-number discrimination. Noise that fluctuates with the envelope of a natural chorus is attractive to female gray treefrogs and, thus, has behavioral salience as a biological signal indicative of a breeding chorus (e.g., Swanson *et al.*, 2007). Such noises, therefore, have potential to confound interpretations of behavioral responses in ways not characteristically considered in studies of dip listening. To avoid these potentially problematic confounds, we used two types of modulated chorus-shaped noises in our experiment. We used

sinusoidally amplitude modulated (SAM) noise because of its widespread use in hearing research and because gray treefrogs experience masking release in SAM noises (Vélez and Bee, 2011). Both unmodulated noise and SAM noise are behaviorally neutral and are neither attractive nor repulsive to gray treefrogs (Vélez and Bee, 2010, 2011). Because of the increasing importance of using biologically relevant natural sounds in hearing research (Lewicki, 2002; Singh and Theunissen, 2003; McDermott and Simoncelli, 2011), chorus amplitude modulated (CAM) noise was used as a second type of modulated noise. CAM noises were designed to have the same modulation spectra as natural choruses (i.e., they fluctuated at rates characteristic of natural choruses), but lacked the discrete temporal fluctuations associated with actual pulsatile calls, which make natural chorus noise attractive to females (Swanson *et al.*, 2007). As part of this study, a control experiment was conducted to confirm that CAM noises were, in fact, behaviorally neutral sounds.

C. The current work

The present study had two objectives. The first objective was to replicate earlier work showing that females prefer longer calls in quiet and that noise reduces this preference. Only upon corroboration of these two findings could we meaningfully achieve our second objective. This second objective was to test the hypothesis that reductions in preferences for longer calls in noise are less severe in temporally modulated noise compared with unmodulated noise. Results consistent with this hypothesis would suggest a role for dip listening in pulse-number discrimination under noisy listening conditions. To test this dip-listening hypothesis, we used two-stimulus choice tests and a phonotaxis paradigm (Gerhardt, 1995) in which subjects responded by choosing one of two alternating calls differing only in pulse number in quiet or in the presence of unmodulated, SAM, or CAM noise. This study successfully replicated previously established patterns of preferences for longer calls and the negative impacts of noise. However, there was no evidence to suggest that pulse-number discrimination was less severely impaired in SAM or CAM noises relative to unmodulated noise. The importance of these negative results is discussed in relation to previous work on dip listening in gray treefrogs and other nonhuman animals.

II. METHODS

A. Subjects

Our procedures for collecting, testing, and handling subjects were approved by the University of Minnesota's Institutional Animal Care and Use Committee (#0809A46721, November, 2008) and were 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 procedures also closely followed those described in detail in previous studies (e.g., Bee and Schwartz, 2009; Bee *et al.*, 2012). Readers are referred to these studies for additional methodological details that are not repeated here. Briefly, between May and June 2009 and

2010, we collected females of the western mitochondrial DNA lineage of Cope's gray treefrog (Ptacek *et al.*, 1994) found in amplexus between 2100 and 0100 h in several ponds and wetlands at the Carver Park Reserve (44° 52' 49.29" N, 93° 43' 3.10" W; Carver County, MN, USA) and the Crow Hassan Park Reserve (45° 11' 18.71" N, 93° 39' 9.05" W; Hennepin County, MN, USA). Subjects were transported to the lab and kept at 2 °C to delay egg deposition prior to testing. On the day of testing, subjects were transferred to an incubator at least 45 min prior to testing to allow their body temperature to reach 20° ± 1 °C. After completing all tests, subjects were released at their location of capture (usually within 48 h of collection). In total, 374 females were collected and tested as part of this study.

B. Apparatus

Phonotaxis tests were conducted inside one of two temperature-controlled (20° ± 2 °C), hemi-anechoic sound chambers (details in Bee and Schwartz, 2009). In quiet conditions, the background noise level in each chamber ranged between 2 and 12 dB SPL (fast rms, flat weighting) in the $\frac{1}{3}$ -octave bands between 500 and 4000 Hz, which span the frequency range of interest in this study. Subjects were tested in a circular arena (2-m diameter) made of 60-cm high walls that were acoustically transparent but visually opaque. The perimeter of the test arena was divided into 24, 15° arcs using short, thin strips of adhesive tape placed on the arena floor (see Fig. 3 in Nityananda and Bee, 2012). Digital acoustic stimuli were broadcast using Adobe Audition v1.5 (Adobe Systems Inc., San Jose, CA) running on a Dell Computer (Optiplex GX620 or GX745; Dell Computer Corp., Round Rock, TX) that was located outside the sound chamber. Sound output was delivered via a M-Audio FireWire 410 multichannel soundcard (M-Audio USA, Irwindale, CA) and amplified using either a Sonamp 1230 (Sonance, San Clemente, CA) or HTD 1235 (Home Theater Direct Inc., Plano, TX) multichannel amplifier. The frequency response of the playback system was flat (±2.5 dB) between 500 and 5000 Hz.

Target signals (described in Sec. II C 1) were broadcast from two A/D/S L210 speakers (Directed Electronics, Inc., Vista, CA) placed on the floor of the sound chamber. Each speaker was centered within one of the 15° arcs around the perimeter, just outside the wall of the arena, and directed toward a release point at the arena's center. The two speakers used to present the two alternative stimuli in a choice test were located 180° apart and separated by 2 m on opposite sides of the arena. Over the course of the experiment, the absolute positions of the speakers around the arena's perimeter were varied between tests of every two to four subjects to eliminate any influence of directional response bias. Noises (described in Sec. II C 2) were broadcast from a Kenwood KFC-1680ie speaker (Kenwood USA Corporation, Long Beach, CA) suspended from the ceiling of the chamber 190 cm above the center of the arena. This overhead speaker created uniform noise levels (±2 dB) across the entire arena floor. Sound levels were measured and calibrated using a Brüel & Kjær type 2250 (Brüel & Kjær North America Inc., Norcross, GA) or a Larson Davis System 824 (Larson Davis,

Depew, NY) sound-level meter. For calibrations, the microphone of the sound-level meter was placed at the approximate position of a subject's head at the release point.

C. Acoustic stimuli

All acoustic stimuli were digitally generated using MATLAB v7.6 (MathWorks, Natick, MA) with a sampling rate of 11 025 Hz and 16-bit resolution.

1. Target signals

Target signals were synthetic models of advertisement calls. Examples are depicted in Fig. 1. They consisted of variable numbers of pulses, but all other acoustic properties were held constant. We constructed stimuli by first

FIG. 1. Spectra of target signals and noises and envelopes of target signals. (a) Frequency spectra (fast-Fourier transformation, 256-point, Hamming window) of the target signals (solid gray line) and chorus-shaped noises (dashed gray line—average of 14 recordings; solid, dotted, dashed, and dot-dashed black lines—average of three recordings each). (b)–(d) Waveforms of synthetic calls with (b) 24 pulses (−2 SD), (c) 32 pulses (∼average), and (d) 40 pulses (+2 SD).

generating a single pulse with spectral and temporal properties close to the temperature-corrected (20 °C) mean values of our study population (Ward *et al.*, 2013). The pulse was 11-ms long and consisted of two harmonically related, phase-locked sinusoids having frequencies (and relative amplitudes) of 1250 Hz (−9 dB) and 2500 Hz (0 dB) [Fig. 1(a)]. The amplitude envelope of the pulse was shaped with a 4-ms inverse exponential rise time and a 7-ms exponential fall time. To create calls, we repeated this pulse after a silent, 11-ms inter-pulse interval (i.e., using a 22-ms pulse period and 50% pulse duty cycle) until the desired number of pulses was reached. Within each call, pulses repeated at a rate of 45.5 pulses/s. The average-length call had 32 pulses. The shorter and longer alternatives had 24, 28, 36, and 40 pulses [Figs. 1(b)–1(d)]. In standard deviation (SD) units, the difference between the pulse numbers in the average call and the shorter and longer alternative calls corresponded to −2 SD, −1 SD, +1 SD, and +2SD, respectively (Bee, 2008; Ward *et al.*, 2013). Hereafter, this SD notation is used to denote the pulse number in the alternative to the average call. We shaped each call envelope using a linear rise over the first 50 ms of the call. During a choice test, each stimulus call repeated with a period of 5 s. The two calls were broadcast 180° out of phase with each other as measured from their temporal midpoints, such that there were equal silent intervals preceding and following each call.

All choice tests were replicated using two different SPLs for the target signal, 85 dB and 79 dB (calibrated at the release point). The 85-dB signal level simulated a naturalistic call amplitude for this species (Gerhardt, 1975). We replicated choice tests at two different signal levels in an attempt to generalize our results and because the preferences of female frogs can be level dependent (Gerhardt and Huber, 2002).

2. Unmodulated, SAM, and CAM noises

We created artificial, chorus-shaped noises [Fig. 1(a)] that had a long-term frequency spectrum modeled after that of gray treefrog choruses recorded between May and July of 2007, 2008, 2009, and 2010 (additional details of recording protocols can be found in Vélez and Bee, 2010, 2011). Our noises were based on average spectra determined from 90-s digital recordings (44 100 Hz sampling rate, 16-bit resolution) of between 3 and 14 different choruses. As illustrated in Fig. 1(a), there was little variation in average spectra stemming from computing averages over different numbers of choruses. During choice tests conducted in noise, the noise was broadcast from the overhead speaker and had a long-term equivalent rms amplitude of 73 dB SPL (LC_{eq}) at the release point. This noise level fell within the range of background noise levels others and we have recorded in gray treefrog choruses (Schwartz *et al.*, 2001; Swanson *et al.*, 2007; M. S. Caldwell and M. A. Bee, unpublished data). In conditions with noise, the 85-dB and 79-dB signal levels corresponded to signal-to-noise ratios (SNRs) of +12 dB and +6 dB, respectively. These SNRs were chosen to be above signal-recognition thresholds based on prior work with this species (Bee and Schwartz, 2009; Vélez and Bee, 2011,

2013; Nityananda and Bee, 2012). Noises that were 6 min in duration were created for behavioral testing.

Unmodulated chorus-shaped noises were created as follows. First, chorus recordings were down-sampled to 11 025 Hz, transformed to the spectral domain (FFT size = 4 194 304), and then the average spectrum across the sampled choruses was computed ($N = 3$ to 14 choruses). Next, this average spectrum was multiplied by the frequency spectrum (FFT size = 4 194 304) of an exemplar of white noise. The resulting frequency spectrum was next band-pass filtered between 850 and 3300 Hz by setting the Fourier coefficients outside this range to zero and then transformed back to the temporal domain with an inverse FFT. In total, we generated eight exemplars of unmodulated, chorus-shaped noise, each one based on a different sample of white noise. Figure 2(a) depicts the amplitude envelopes of unmodulated chorus-shaped noise, which exhibited only the random level fluctuations inherent in band-limited noise.

SAM noises were created by multiplying an unmodulated chorus-shaped noise by a sine wave using the equation

$$n(t) = A[1 + m\sin(2\pi f_m t + \rho)]c(t), \quad (1)$$

where $n(t)$ is the resulting SAM noise, A is a scaling factor, m is the modulation depth (1.0 in all cases), f_m is the modulation frequency, ρ is the starting phase of the sine wave, $c(t)$ is the unmodulated chorus-shaped noise, and t is time in seconds. We used modulation frequencies of 0.625 and 45 Hz to create slow [SAM(slow)] and fast [SAM(fast)] fluctuating noises, respectively [Figs. 2(b) and 2(c)]. Rates of 0.625 and 45 Hz were chosen because they were within the most prominent peaks in modulation spectra of recordings of gray treefrog choruses (Vélez and Bee, 2010). Four exemplars of each SAM noise were created, each using a different exemplar of unmodulated noise. Each exemplar also had a different starting phase for the sinusoidal modulator (0°, 90°, 180°, and 270°). Equivalent numbers of subjects were tested with each exemplar.

We created CAM noises using four exemplars of unmodulated, chorus-shaped noise, each with a frequency spectrum computed by averaging the spectra of three different chorus recordings. The first step in creating a CAM noise was extracting the Hilbert envelope of each of the three chorus recordings used to generate the frequency spectrum of the corresponding unmodulated noise exemplar. We then transformed these three envelopes to the frequency domain (FFT size = 4 194 304) and calculated their average temporal-modulation spectrum [Fig. 3(a)]. For each exemplar of unmodulated noise used, we created a CAM noise with both the slow (≤ 2 Hz) and fast (~ 40 –50 Hz) rates of amplitude modulation typical of natural choruses [CAM(both)] as follows. First, the average modulation spectrum was low-pass filtered with a cutoff frequency of 80 Hz. Next we back transformed the resulting spectrum to the time domain using an inverse FFT. Finally, we imposed the resulting envelope on the original exemplar of unmodulated noise that had a frequency spectrum determined from the same three chorus recordings used to generate the average

FIG. 2. Envelope fluctuations of different noises. Shown here are 3.2-s segments of the amplitude envelopes of one exemplar each of the (a) unmodulated, (b) SAM(slow), (c) SAM(fast), (d) CAM(both), (e) CAM(slow), and (f) CAM(fast) noises used in this study.

modulation spectrum. Hence, each resulting CAM(both) noise exemplar [Figs. 2(d) and 3(b)] had an average frequency spectrum and an average modulation spectrum similar to those of the three original chorus recordings used to create it. Similar procedures were followed to generate exemplars of CAM(slow) noise and CAM(fast) noise, which were restricted to have either the slower or the faster fluctuations typical of natural choruses (Figs. 2 and 3). For CAM(slow) noises [Figs. 2(e) and 3(c)] the average

modulation spectrum was low-pass filtered with a cutoff frequency of 15 Hz (instead of 80 Hz), and for CAM (fast) noises [Figs. 2(f) and 3(d)] the average modulation spectrum was band-pass filtered between 15 and 80 Hz. For comparison purposes, the modulation spectrum of an unmodulated noise is depicted in Fig. 3(e). In total, we generated four different exemplars of CAM(both), CAM(slow), and CAM(fast) noise using the same four exemplars of unmodulated noise. Each exemplar was used to test an equivalent number of subjects.

D. Behavioral testing

1. Choice tests

We conducted four, two-stimulus choice tests (Gerhardt, 1995) in which subjects were presented with the average call and either a shorter (-2 SD or -1 SD) or longer ($+1$ SD or $+2$ SD) alternative call. All four tests were replicated at both signal levels (85 dB and 79 dB) in quiet and in each of six different noise conditions [unmodulated, SAM(slow), SAM(fast), CAM(both), CAM(slow), and CAM(fast)]. All together, our experimental design yielded a total of 56 tests (4 tests $\times$ 2 signal levels $\times$ 7 conditions). Each test had a sample size of 40 subjects. Separate groups of subjects were tested at each signal level (i.e., signal level was a between-subjects factor). Within each signal level, however, individual subjects were tested in a variable number of different tests (median = 5; mode = 5; maximum = 13). Each subject contributed only one response per test and was given a time-out of at least 5 min between consecutive tests. Previous studies indicate female gray treefrogs do not exhibit directional biases or carry-over effects as a result of repeated testing in this manner (Gerhardt *et al.*, 2000). The combination of choice tests, signal level, and noise conditions that each subject experienced was determined randomly.

To begin each phonotaxis trial, a single subject was placed in an acoustically transparent holding cage located at the release point in the center of the arena and allowed to sit in quiet during a 1-min acclimatization period. In trials conducted with noise, we began broadcasts of the noise immediately after the end of the acclimatization period. Presentations of alternating signals commenced 30 s after the noise onset. For trials conducted in quiet, subjects experienced an additional 30 s of quiet after the end of the acclimatization period prior to onset of the alternating signals. An experimenter positioned outside of the sound chamber remotely released the subject using a rope and pulley system to lift the lid of the holding cage after three repetitions of the two alternating calls. We scored a choice as occurring when subjects exhibited stereotyped phonotaxis behavior (e.g., Rheinlaender and Klump, 1988) and touched the wall of the test arena inside a 15° arc centered on one of the two signal speakers. Trials were conducted under infrared (IR) illumination. Choices were scored in real time using a video monitor located outside the chamber and an IR-sensitive video camera (Panasonic WV-BP334, Panasonic Corporation of North America, Secaucus, NJ) mounted from the ceiling of the chamber above the center of the test arena. We have

FIG. 3. Modulation spectra depicting the rates of envelope fluctuations in natural and artificial chorus noises. Shown here are the modulation spectra for (a) three separate natural choruses (averaged over 90-s recordings for each) and (b)-(e) one exemplar each of the (b) CAM(both), (c) CAM(slow), (d) CAM(fast) chorus-shaped noises and (e) an unmodulated chorus-shaped noise.

recently confirmed real-time scoring to be as accurate as blind post-test scoring from videos (Bee *et al.*, 2012). Subjects were given up to 5 min to make a choice. Typical response latencies are less than 2 min. Subjects that failed to make a choice within 5 min ($N = 28$) were considered non-responsive and were excluded from the final dataset, which included 2240 choices made by 326 subjects.

2. Control experiment

To assess whether CAM noises were behaviorally neutral, we conducted an experiment in which these sounds were presented as potential target signals. Our protocol used phonotaxis tests generally similar to those described for choice tests (Sec. IID 1) and in parallel control tests conducted with unmodulated and SAM noises in previous studies (Vélez and Bee, 2010, 2011, 2013; Vélez *et al.*, 2012). Readers are referred to those studies for details not reported here. Briefly, using a within-subjects design, subjects ($N = 20$) were tested in a series of single-stimulus phonotaxis tests comprising three experimental treatments. In the experimental treatments, each CAM noise [CAM(both), CAM(slow), and CAM(fast)] was broadcast continuously from a speaker on the floor, just outside the wall of the arena, from the same positions from which we broadcast one of the target signals in choice tests. We calibrated the long-term rms amplitude of all CAM noise stimuli to 73 dB SPL (LC_{eq}) at the release point. After a 1-min acclimatization period, followed by 30 s of noise, subjects were released and given up to 5 min to touch the wall of the test arena. The response variable was the 15° arc in which subjects first touched the wall of the arena relative to the arc in which the speaker was located, which was designated as 0° .

E. Statistical analysis

The lme4 package (Bates *et al.*, 2011) in R version 3.0.0 (R Core Team, 2013) was used to construct Generalized Linear Mixed Models (GLMM) to analyze the responses of subjects across different choice tests. The binary response variable was the call chosen by each subject in each test. A choice of the relatively longer call was assigned a value of 1 and a choice of the shorter call was assigned a value of 0. A logit function was used to link the probability of choosing longer calls to the linear predictors in all models. Subject was included as a random factor in all models. Specifically, we computed logistic mixed-effect models in which, for the i th subject ($i = 1, 2, \dots, N$) and the j th observation nested within subject ($j = 1, 2, \dots, n_i$), the linear predictor was augmented with a single random effect for subject i according to the following equations,

$$P(Y_{ij} = 1 | x_{ij}, \alpha_{ij}) = p_{ij}, \quad (2)$$

$$\log\left(\frac{p_{ij}}{1 - p_{ij}}\right) = x'_{ij}\beta + \alpha_i, \quad (3)$$

where Y_{ij} 's are the binary responses, x_{ij} 's are the design vectors for the fixed effects constructed from the predictors, β is the vector of coefficients for the fixed effects and α_i 's are the subject-specific random effects with $\alpha_i \sim N(0, \sigma^2)$. Noise exemplar (1-8) was included as a blocking variable in all models. For tests conducted in quiet, we assigned the same exemplar number used to test the same subject in noise.

An initial model assessed whether this study was able to replicate earlier work on pulse-number discrimination in quiet and in noise (Sec. IA). Specifically, this model assessed whether subjects preferentially chose longer calls and whether

the presence of noise reduced the probability that subjects would choose longer calls compared with quiet conditions. A full model was constructed that included the following categorical fixed effects and their two-way and three-way interactions: noise (present versus absent), signal level (85 dB versus 79 dB), and pulse number in the alternative to the average call (± 1 SD or ± 2 SD). We assessed whether the probability that subjects chose longer calls was greater than expected by chance (0.5) using Z-tests to determine whether the log odds ratio was significantly greater than zero. A significant main effect of noise was predicted if noise reduced preferences for longer calls. To the extent that subjects were either more or less selective depending on signal level, we predicted a significant difference between responses measured at the 85-dB and 79-dB playback levels. There was no *a priori* directional expectation for the effect of signal level. We expected nonlinearity in preferences for longer calls to be reflected in a significant effect of pulse number in the alternative, with a relatively higher probability of choosing longer calls when shorter-than-average alternatives were used (-2 SD and -1 SD). If the effects of noise depended on either the SNR or the number of pulses in the alternative, we expected to find either a noise $\times$ signal level interaction or a noise $\times$ pulse number interaction, respectively.

In a second model, we tested the main hypothesis of interest, namely, that the temporal envelope of the noise influenced the probability that subjects would choose longer calls. The full model included the following categorical fixed effects and their two-way and three-way interactions: noise type [unmodulated, SAM(slow), SAM(fast), CAM(both), CAM(slow), or CAM(fast)], signal level (85 dB versus 79 dB), and pulse number in the alternative to the average call (± 1 SD or ± 2 SD). According to a dip listening hypothesis, our main prediction was that subjects should be more likely to choose longer calls in modulated noise than in unmodulated noise. This should be reflected in a significant effect of noise type. We expected to find a significant noise type $\times$ signal level interaction if the effect of noise type depended on the SNR. A significant interaction between noise type and pulse number in the alternative would indicate that the effects of noise type varied depending on the alternatives presented to subjects. Our expectations for the effects of signal level and pulse number in the alternative were the same as described for the first model.

Each of the two full models was compared against a corresponding additive model that included only main effects using the Akaike Information Criterion (AIC) and a χ^2 test. If there was no significant difference between the full and additive models, we based our interpretation of results on the reduced model. The significance of separate effects in the final model was assessed using type II Wald χ^2 tests.

Results from the control test were analyzed using circular statistics (Batschelet, 1981) computed with the Oriana 3.0 software package. The null hypothesis that the angles at which subjects first touched the wall of the test arena were uniformly distributed was tested using Rayleigh tests. If CAM noises either attracted or repelled females, we predicted orientation angles should be directed toward or away from the stimulus, respectively. A significance criterion of $\alpha = 0.05$ was used for all statistical tests.

III. RESULTS

A. Pulse-number discrimination in quiet and in noise

Figure 4(a) depicts responses in quiet and in noise as a function of the number of pulses in the alternative to the average call. The full model (AIC = 2944.66) did not provide a significantly better fit to these data than the additive model (AIC = 2936.90) ($\chi^2 = 12.24$, $df = 10$, $P = 0.2690$). In the additive model (see Table I), there were significant effects of noise, signal level, and pulse number in the alternative. Noise exemplar was not significant. The probability that subjects chose longer calls was significantly greater than 0.5 in both quiet ($Z = 6.08$, $P < 0.0001$) and in noise ($Z = 8.81$, $P < 0.0001$). As illustrated in Fig. 4(a), however, subjects were significantly less likely to choose the longer alternative in the presence of noise compared with quiet (Tables I and II). Subjects were also less likely to choose longer calls at the 79-dB signal level compared with the 85-dB signal level [Fig. 4(a); Tables I and II]. The significant effect of pulse number in the alternative indicates that preferences varied according to the number of pulses in the choices offered to subjects. As illustrated in Fig. 4(a), and as confirmed in a contrast analysis ($Z = 5.38$, $P < 0.0001$), subjects were more likely to choose longer calls when the short alternative was shorter than average (-2 SD and -1 SD) than when the average call itself was the short alternative ($+1$ SD and $+2$ SD).

B. Pulse-number discrimination in modulated and unmodulated noise

Figure 4(b) depicts responses in the six different noise conditions for the 85-dB and 79-dB signal levels. The full model (AIC = 2576.00) did not provide a significantly better fit to these data than the additive model (AIC = 2551.72) ($\chi^2 = 51.73$, $df = 38$, $P = 0.0679$). In the additive model, the effect of noise type was not significant (Tables III and IV). There were significant effects of signal level and pulse number in the alternative (Tables III and IV). The general patterns in responses due to these two significant effects [Fig. 4(b)] were generally similar to those described for responses collapsed over all noise types [Fig. 4(a)].

C. Control experiment

Table V summarizes the results of the control experiment. None of the distributions of orientation angles in response to CAM noises were different from uniform (all $Z_s < 0.28$; all $P_s > 0.7$). The lengths of mean vectors (r) were less than or equal to 0.15 and circular standard deviations were greater than 100° .

IV. DISCUSSION

Three results merit discussion before considering the main finding of this study related to pulse-number discrimination in modulated versus unmodulated noise. Two of these results corroborate earlier work on gray treefrogs (Gerhardt, 1994; Gerhardt *et al.*, 1996; Gerhardt *et al.*, 2000; Schwartz *et al.*, 2001; Bee, 2008; Ward *et al.*, 2013). First, when given a choice between two calls differing only in pulse number,

FIG. 4. Results from two-stimulus choice tests of pulse-number discrimination. Shown here are the proportions ($\hat{p} \pm$ exact 95% binomial confidence intervals) of subjects choosing the longer call as a function of the pulse number in the alternative to the average call. (a) Results in quiet and in noise (collapsed over all noise conditions) at the 85-dB and 79-dB signal levels. (b) Results for unmodulated noise and five types of modulated noise at the 85-dB and 79-dB signal levels. Horizontal dashed lines depict chance expectation (0.5) in a two-stimulus choice test.

subjects preferred longer calls. Moreover, preferences in favor of longer calls were nonlinear, in that they were more pronounced when the alternative was shorter than average. Second, subjects were less likely to choose longer calls in the presence of noise compared with quiet conditions. Together, these two results confirm earlier work and suggest that noise interferes with the expression of preferences for longer calls. Finally, CAM noises were neither attractive nor repulsive when presented as potential target signals. This third result indicates that CAM noises were behaviorally neutral stimuli and, thus, similar in this regard to the unmodulated and SAM noises (Vélez and Bee, 2010, 2011).

Our replication of earlier work demonstrating preferences for longer calls and reduced preferences in noise was critically important, because it provided the necessary framework for evaluating evidence related to our main hypothesis of interest. Recall that according to the dip listening

hypothesis, subjects should be more likely to choose longer calls in modulated noise compared with unmodulated noise. The evidence was not consistent with this prediction. There were no significant differences in the proportions of subjects

TABLE II. Parameter estimates from the additive GLMM model of pulse-number discrimination in quiet and in noise.^a

Fixed effects	Estimate	Std. Err.	Z	P
Intercept	1.31	0.21	6.36	<0.0001
Noise (present)	-0.30	0.13	-2.28	0.0225
Signal level (79 dB)	-0.32	0.09	-3.44	0.0006
Pulse number in alternative (-1 SD)	-0.48	0.13	-3.66	0.0002
Pulse number in alternative (+1 SD)	-0.89	0.13	-6.86	<0.0001
Pulse number in alternative (+2 SD)	-0.56	0.13	-4.30	<0.0001
Noise exemplar (2)	-0.08	0.20	-0.42	0.6772
Noise exemplar (3)	-0.14	0.20	-0.71	0.4787
Noise exemplar (4)	0.30	0.21	1.44	0.1503
Noise exemplar (5)	0.18	0.19	0.94	0.3482
Noise exemplar (6)	0.06	0.19	0.33	0.7431
Noise exemplar (7)	0.19	0.19	1.00	0.3154
Noise exemplar (8)	0.13	0.19	0.69	0.4906
Random effects	Variance	Std. Dev.		
Subject	0.07	0.27		

^aThe reference levels for each fixed effect were as follows: Noise (absent), Signal level (85 dB), Pulse number in alternative (-2 SD), and Noise exemplar (1).

TABLE I. Analysis of deviance table for the additive GLMM model of pulse-number discrimination in quiet and in noise.

Effect	χ^2	df	P
Noise (present or absent)	5.21	1	0.0225
Signal level (85 dB or 79 dB)	11.81	1	0.0006
Pulse number in alternative (± 1 or ± 2 SD)	47.53	3	<0.0001
Noise exemplar (1-8)	7.86	7	0.3449

TABLE III. Analysis of deviance table for the additive GLMM model of pulse-number discrimination in modulated and unmodulated noises.

Effect	χ^2	df	P
Noise type (unmodulated, SAM(slow), SAM(fast), CAM(both), CAM(slow), and CAM(fast))	4.78	5	0.4439
Signal level (85 dB or 79 dB)	8.54	1	0.0035
Pulse number in alternative (± 1 or ± 2 SD)	34.92	3	<0.0001
Noise exemplar (1-8)	7.98	7	0.3340

choosing longer calls between unmodulated noise and five types of modulated noise. Hence, the new finding from the present study is that pulse-number discrimination was not improved in modulated noise conditions as a result of dip listening.

Clearly some caution is necessary when interpreting negative results consistent with a null hypothesis of “no difference.” We suggest the present study had a reasonably good chance of detecting a difference in pulse-number discrimination in unmodulated versus modulated noise if this difference were not exceedingly small. This suggestion is based, in part, on the large size of the present study relative to other studies. In total, we tested the preferences of 326 subjects in 2240 choice tests, with a minimum of 80 subjects and 320 choice tests in each particular noise condition. These sample sizes are considerably larger than those used to demonstrate dip listening by measuring signal-recognition thresholds (e.g., $N = 240$ frogs, with 24 frogs per noise condition in Véléz and Bee, 2011). In addition, this study used a large factorial design to test five different types of modulated noise, two signal levels, and four levels of difference in pulse

TABLE IV. Parameter estimates from the additive GLMM model of pulse-number discrimination in modulated and unmodulated noise.^a

Fixed effects	Estimate	Std. Err.	Z	P
Intercept	1.06	0.23	4.57	<0.0001
Noise [SAM(slow)]	-0.01	0.20	-0.04	0.9650
Noise [SAM(fast)]	-0.26	0.20	-1.30	0.1944
Noise [CAM(both)]	-0.22	0.20	-1.07	0.2859
Noise [CAM(slow)]	-0.03	0.20	-0.13	0.8934
Noise [CAM(fast)]	-0.07	0.20	-0.33	0.7380
Signal level (79 dB)	-0.28	0.10	-2.92	0.0035
Pulse number in alternative (-1 SD)	-0.47	0.14	-3.40	0.0007
Pulse number in alternative (+1 SD)	-0.81	0.14	-5.89	<0.0001
Pulse number in alternative (+2 SD)	-0.49	0.14	-3.54	0.0004
Noise exemplar (2)	-0.08	0.21	-0.40	0.6872
Noise exemplar (3)	-0.16	0.21	-0.79	0.4293
Noise exemplar (4)	0.32	0.21	1.50	0.1349
Noise exemplar (5)	0.15	0.29	0.51	0.6079
Noise exemplar (6)	-0.02	0.29	-0.08	0.9387
Noise exemplar (7)	0.21	0.29	0.73	0.4637
Noise exemplar (8)	0.05	0.29	0.19	0.8520
Random effects	Variance	Std. Dev.		
Subject	0.01	0.11		

^aThe reference levels for each fixed effect were as follows: Noise type (unmodulated), Signal level (85 dB), Pulse number in alternative (-2 SD), and Noise exemplar (1).

TABLE V. Results of circular statistical analyses of orientation in response to CAM noise.^a

Treatment	N	Mean vector (°)	Length of mean vector (r)	Circular standard deviation (°)	Rayleigh test (Z)	Rayleigh test (P)
CAM(both)	15	255.8	0.13	116.5	0.241	0.79
CAM(slow)	15	252.3	0.07	132.9	0.069	0.94
CAM(fast)	13	45.2	0.15	112.5	0.275	0.77

^aThe column headed N reports the number of subjects (out of 20) that reached the wall of the arena within 5 min of stimulus presentation.

number. Hence, the negative results reported here generalize across a fairly broad range of stimulus conditions. Nevertheless, we remain open to the possibility that some benefit of dip listening to pulse-number discrimination might be found in a future study over a very limited range of lower SNRs (e.g., +3 dB) between those required to elicit behavioral responses in noise (e.g., -3 to 0 dB) and the suprathreshold SNRs tested here (e.g., +6 to +12 dB). Schwartz *et al.* (2013) failed to find evidence that signal-recognition thresholds themselves varied according to pulse number in either modulated or unmodulated noise.

Taking our main finding at face value, our results stand in stark contrast with those from studies of dip listening that measured perceptual performance in terms of signal-recognition thresholds. Both Cope’s gray treefrogs (Véléz and Bee, 2011, 2013) and eastern gray treefrogs (Schwartz *et al.*, 2013) have lower signal-recognition thresholds in some modulated noises compared with unmodulated noises. Females of Cope’s gray treefrog, for example, recognize male advertisement calls at significantly lower SNRs when the dips in slowly fluctuating SAM noises are long enough to include a minimum of six to nine consecutive pulses of the call (Véléz and Bee, 2011). Based on these earlier results, and the finding that chorus noise is dominated by low-frequency modulations (≤ 2 Hz), we had predicted that a male producing relatively longer calls should have a better chance of placing more pulses of his calls into the dips in chorus background noise and, therefore, a greater chance to attract females (Véléz and Bee, 2011). Even if there were no advantage of longer calls in terms of improved signal-recognition thresholds (e.g., Schwartz *et al.*, 2013), males that produced relatively longer calls might still benefit from doing so if dip listening resulted in females preferentially choosing relatively longer calls. Results from the present study, however, do not support this earlier prediction.

The contrast between our current results and previous work showing a benefit of modulated noise is important because it highlights the potential limits of dip listening in frogs. In Cope’s gray treefrogs, the ability to catch acoustic glimpses of signals during dips in background noise is beneficial in terms of recognizing conspecific signals (Véléz and Bee, 2011, 2013). Call recognition in this species is mediated by perception of the signal’s species-specific pulse rate (Schul and Bush, 2002). Similar benefits, however, do not appear to extend to intraspecific call discrimination along a continuum of individual differences in pulse number. It will be important

in future work on gray treefrogs to investigate dip listening in the context of call discrimination along other biologically relevant continua, such as individual differences in call spectrum (Schrode *et al.*, 2012) and call rate (Ward *et al.*, 2013).

Few previous studies have assessed dip listening in the context of acoustic communication in animals other than humans and frogs. Ronacher and Hoffmann (2003) investigated the ability of male bow-winged grasshoppers (*Chorthippus biguttulus*) to recognize a stridulatory signal produced by females in the presence of modulated and unmodulated noise. Interestingly, they found that males experienced a release from masking in the presence of slowly fluctuating noises when the target signal was long (12 subunits), but not when it was short (4 subunits). This result in grasshoppers, thus, contrasts with those reported here and elsewhere (Schwartz *et al.*, 2013) for gray treefrogs. Behavioral studies of fish (Fay, 2011), birds (e.g., Klump and Langemann, 1995; Jensen, 2007), and dolphins (Branstetter and Finneran, 2008; Trickey *et al.*, 2010; Branstetter *et al.*, 2013) have shown that thresholds for detecting simple acoustic stimuli (e.g., tones and narrowband noises) are lower in the presence of temporally fluctuating noises compared with unmodulated noises. The importance of these findings in relation to perceiving acoustic communication signals remains to be determined.

V. CONCLUSIONS

Results from earlier work on Cope's gray treefrog and other frog species support the emerging view that frogs exploit some of the same perceptual cues as humans for solving a communication problem analogous to the human cocktail party problem (reviewed in Bee, 2012; Vélez *et al.*, 2014). The negative results reported here are important in the broader context framed by this earlier work on frogs because they demonstrate that the perceptual benefits of exploiting some cues also have limits. Based on the currently available evidence, we provisionally conclude that dip listening is beneficial to gray treefrogs in the context of recognizing conspecific calls based on the species-specific pulse rate, but not in the context of intraspecific discrimination among individuals whose calls vary in pulse number.

ACKNOWLEDGMENTS

This work was supported by Grant No. R01 DC 009582 from the National Institutes of Health. The authors thank M. Linck and the Three Rivers Park District for access to study sites, D. Pereira and the Minnesota Department of Natural Resource for permission to collect frogs, N. Lee for feedback on an earlier version of the manuscript, and S. Tekmen and numerous undergraduate research assistants for help collecting and testing frogs.

Arak, A. (1983). "Sexual selection by male-male competition in natterjack toad choruses," *Nature* **306**, 261–262.
 Bacon, S. P., Opie, J. M., and Montoya, D. Y. (1998). "The effects of hearing loss and noise masking on the masking release for speech in temporally complex backgrounds," *J. Speech Lang. Hear. Res.* **41**, 549–563.
 Bates, D., Maechler, M., and Bolker, B. (2011). "lme4: Linear mixed-effects models using Eigen and R syntax," <http://R-Forge.R-project.org/projects/lme4/> (Last viewed 12 July 2013).

Batschelet, E. (1981). *Circular Statistics in Biology* (Academic Press, London), 371 pp.
 Bee, M. A. (2008). "Parallel female preferences for call duration in a diploid ancestor of an allotetraploid treefrog," *Anim. Behav.* **76**, 845–853.
 Bee, M. A. (2012). "Sound source perception in anuran amphibians," *Curr. Opin. Neurobiol.* **22**, 301–310.
 Bee, M. A., and Micheyl, C. (2008). "The cocktail party problem: What is it? How can it be solved? And why should animal behaviorists study it?," *J. Comp. Psych.* **122**, 235–251.
 Bee, M. A., and Schwartz, J. J. (2009). "Behavioral measures of signal recognition thresholds in frogs in the presence and absence of chorus-shaped noise," *J. Acoust. Soc. Am.* **126**, 2788–2801.
 Bee, M. A., Vélez, A., and Forester, J. D. (2012). "Sound level discrimination by gray treefrogs in the presence and absence of chorus-shaped noise," *J. Acoust. Soc. Am.* **131**, 4188–4195.
 Branstetter, B. K., and Finneran, J. J. (2008). "Comodulation masking release in bottlenose dolphins (*Tursiops truncatus*)," *J. Acoust. Soc. Am.* **124**, 625–633.
 Branstetter, B. K., Trickey, J. S., Bakhtiari, K., Black, A., Aihara, H., and Finneran, J. J. (2013). "Auditory masking patterns in bottlenose dolphins (*Tursiops truncatus*) with natural, anthropogenic, and synthesized noise," *J. Acoust. Soc. Am.* **133**, 1811–1818.
 Buus, S. (1985). "Release from masking caused by envelope fluctuations," *J. Acoust. Soc. Am.* **78**, 1958–1965.
 Cherry, E. C. (1953). "Some experiments on the recognition of speech, with one and with two ears," *J. Acoust. Soc. Am.* **25**, 975–979.
 Fay, R. R. (2011). "Signal-to-noise ratio for source determination and for a comodulated masker in goldfish, *Carassius auratus*," *J. Acoust. Soc. Am.* **129**, 3367–3372.
 Festen, J. M., and Plomp, R. (1990). "Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing," *J. Acoust. Soc. Am.* **88**, 1725–1736.
 Füllgrabe, C., Berthommier, F., and Lorenzi, C. (2006). "Masking release for consonant features in temporally fluctuating background noise," *Hearing Res.* **211**, 74–84.
 Gerhardt, H. C. (1975). "Sound pressure levels and radiation patterns of vocalizations of some North American frogs and toads," *J. Comp. Physiol.* **102**, 1–12.
 Gerhardt, H. C. (1994). "Reproductive character displacement of female mate choice in the gray treefrog *Hyla chrysoscelis*," *Anim. Behav.* **47**, 959–969.
 Gerhardt, H. C. (1995). "Phonotaxis in female frogs and toads: Execution and design of experiments," in *Methods in Comparative Psychoacoustics*, edited by G. M. Klump, R. J. Dooling, R. R. Fay, and W. C. Stebbins (Birkhäuser Verlag, Basel, Switzerland), pp. 209–220.
 Gerhardt, H. C. (2001). "Acoustic communication in two groups of closely related treefrogs," *Adv. Stud. Behav.* **30**, 99–167.
 Gerhardt, H. C., Dyson, M. L., and Tanner, S. D. (1996). "Dynamic properties of the advertisement calls of gray tree frogs: Patterns of variability and female choice," *Behav. Ecol.* **7**, 7–18.
 Gerhardt, H. C., and Huber, F. (2002). *Acoustic Communication in Insects and Anurans: Common Problems and Diverse Solutions* (Chicago University Press, Chicago), 531 pp.
 Gerhardt, H. C., Tanner, S. D., Corrigan, C. M., and Walton, H. C. (2000). "Female preference functions based on call duration in the gray tree frog (*Hyla versicolor*)," *Behav. Ecol.* **11**, 663–669.
 Gustafsson, H. A., and Arlinger, S. D. (1994). "Masking of speech by amplitude-modulated noise," *J. Acoust. Soc. Am.* **95**, 518–529.
 Jensen, K. K. (2007). "Comodulation detection differences in the hooded crow (*Corvus corone cornix*), with direct comparison to human subjects," *J. Acoust. Soc. Am.* **121**, 1783–1789.
 Jin, S. H., and Nelson, P. B. (2006). "Speech perception in gated noise: The effects of temporal resolution," *J. Acoust. Soc. Am.* **119**, 3097–3108.
 Klump, G. M., and Langemann, U. (1995). "Comodulation masking release in a songbird," *Hearing Res.* **87**, 157–164.
 Lewicki, M. S. (2002). "Efficient coding of natural sounds," *Nat. Neurosci.* **5**, 356–363.
 McDermott, J. H., and Simoncelli, E. P. (2011). "Sound texture perception via statistics of the auditory periphery: Evidence from sound synthesis," *Neuron* **71**, 926–940.
 Narins, P. M. (1982). "Effects of masking noise on evoked calling in the Puerto Rican coqui (Anura, Leptodactylidae)," *J. Comp. Physiol.* **147**, 439–446.

- Nityananda, V., and Bee, M. A. (2012). "Spatial release from masking in a free-field source identification task by gray treefrogs," *Hearing Res.* **285**, 86–97.
- Ptacek, M. B., Gerhardt, H. C., and Sage, R. D. (1994). "Speciation by polyploidy in treefrogs: Multiple origins of the tetraploid, *Hyla versicolor*," *Evolution* **48**, 898–908.
- R Core Team (2013). "R: A language and environment for statistical computing," R Foundation for Statistical Computing, Vienna, Austria, <http://www.R-project.org/> (Last viewed 12 July 2013).
- Rheinlaender, J., and Klump, G. M. (1988). "Behavioral aspects of sound localization," in *The Evolution of the Amphibian Auditory System*, edited by B. Fritsch, M. J. Ryan, W. Wilczynski, and T. Hetherington (Wiley and Sons, New York), pp. 297–305.
- Ronacher, B., and Hoffmann, C. (2003). "Influence of amplitude modulated noise on the recognition of communication signals in the grasshopper *Chorthippus biguttulus*," *J. Comp. Physiol. A* **189**, 419–425.
- Schrode, K., Ward, J. L., Vélez, A., and Bee, M. A. (2012). "Female preferences for spectral call properties in the western genetic lineage of Cope's gray treefrog (*Hyla chrysoscelis*)," *Behav. Ecol. Sociobiol.* **66**, 1595–1606.
- Schul, J., and Bush, S. L. (2002). "Non-parallel coevolution of sender and receiver in the acoustic communication system of treefrogs," *Proc. R. Soc. B* **269**, 1847–1852.
- Schwartz, J. J., Buchanan, B. W., and Gerhardt, H. C. (2001). "Female mate choice in the gray treefrog (*Hyla versicolor*) in three experimental environments," *Behav. Ecol. Sociobiol.* **49**, 443–455.
- Schwartz, J. J., Crimarco, N. C., Bregman, Y., and Umeoji, K. (2013). "An investigation of the functional significance of responses of the gray treefrog (*Hyla versicolor*) to chorus noise," *J. Herpetol.* **47**, 354–360.
- Singh, N. C., and Theunissen, F. E. (2003). "Modulation spectra of natural sounds and ethological theories of auditory processing," *J. Acoust. Soc. Am.* **114**, 3394–3411.
- Swanson, E. M., Tekmen, S. M., and Bee, M. A. (2007). "Do female anurans exploit inadvertent social information to locate breeding aggregations?," *Can. J. Zool.* **85**, 921–932.
- Trickey, J. S., Branstetter, B. K., and Finneran, J. J. (2010). "Auditory masking of a 10 kHz tone with environmental, comodulated, and Gaussian noise in bottlenose dolphins (*Tursiops truncatus*)," *J. Acoust. Soc. Am.* **128**, 3799–3804.
- Vélez, A., and Bee, M. A. (2010). "Signal recognition by frogs in the presence of temporally fluctuating chorus-shaped noise," *Behav. Ecol. Sociobiol.* **64**, 1695–1709.
- Vélez, A., and Bee, M. A. (2011). "Dip listening and the cocktail party problem in gray treefrogs: Signal recognition in temporally fluctuating noise," *Anim. Behav.* **82**, 1319–1327.
- Vélez, A., and Bee, M. A. (2013). "Signal recognition by green treefrogs (*Hyla cinerea*) and Cope's gray treefrogs (*Hyla chrysoscelis*) in naturally fluctuating noise," *J. Comp. Psych.* **127**, 166–178.
- Vélez, A., Schwartz, J. J., and Bee, M. A. (2014). "Anuran acoustic signal perception in noisy environments," in *Animal Communication and Noise*, edited by H. Brumm (Springer, New York) (in press).
- Ward, J. L., Love, E. K., Vélez, A., Buerkle, N. P., O'Bryan, L. R., and Bee, M. A. (2013). "Multitasking males and multiplicative females: Dynamic signalling and receiver preferences in Cope's gray treefrog (*Hyla chrysoscelis*)," *Anim. Behav.* **86**, 231–243.
- Welch, A. M., Semlitsch, R. D., and Gerhardt, H. C. (1998). "Call duration as an indicator of genetic quality in male gray treefrogs," *Science* **280**, 1928–1930.