ORIGINAL PAPER

Dip listening or modulation masking? Call recognition by green treefrogs (*Hyla cinerea*) in temporally fluctuating noise

Alejandro Vélez · Gerlinde Höbel · Noah M. Gordon · Mark A. Bee

Abstract Despite the importance of perceptually separating signals from background noise, we still know little about how nonhuman animals solve this problem. Dip listening, an ability to catch meaningful 'acoustic glimpses' of a target signal when fluctuating background noise levels momentarily drop, constitutes one possible solution. Amplitude-modulated noises, however, can sometimes impair signal recognition through a process known as modulation masking. We asked whether fluctuating noise simulating a breeding chorus affects the ability of female green treefrogs (*Hyla cinerea*) to recognize male advertisement calls. Our analysis of recordings of the sounds of green treefrog choruses reveal that their levels fluctuate primarily at rates below 10 Hz. In laboratory phonotaxis tests, we found no evidence for dip listening or modulation masking. Mean signal recognition thresholds in the presence of fluctuating chorus-like noises were never statistically different from those in the presence of a non-fluctuating control. An analysis of statistical effects sizes indicates that masker fluctuation rates, and the presence versus absence of fluctuations, had negligible effects on subject behavior. Together, our results suggest that females listening in natural settings should receive no benefits, nor experience any additional constraints, as a result of level fluctuations in the soundscape of green treefrog choruses.

Keywords Cocktail party problem · Dip listening · Modulation masking · Green treefrog · *Hyla cinerea*

A. Vélez (✉) · M. A. Bee
Department of Ecology, Evolution, and Behavior, University of Minnesota, 100 Ecology, 1987 Upper Buford Circle, St. Paul, MN 55108, USA
e-mail: velez011@umn.edu; avelezme@purdue.edu

Present Address:
A. Vélez
Department of Biological Sciences, Purdue University, 915 West State Street, West Lafayette, IN 47907, USA

G. Höbel
Department of Biological Sciences, University of Wisconsin, 3209 N. Maryland Ave., Milwaukee, WI 53201, USA

N. M. Gordon
Department of Biology, University of Evansville, 219 Koch Center, 1800 Lincoln Ave., Evansville, IN 47722, USA

Introduction

In natural environments, acoustic communication is often compromised because abiotic noise (e.g., wind, rushing water, rustling leaves) and biotic noise (e.g., other signaling animals) impair signal detection, recognition, and discrimination (Klump 1996; Brumm and Slabbekoorn 2005, Langemann and Klump 2005; Arch and Narins 2008; Schwartz and Freeberg 2008). Noise generated by other conspecific signals is an important source of auditory masking in animal social aggregations. Since the frequency spectrum of this type of biotic noise overlaps with that of the target signal, frequency filtering provides little gain in signal-to-noise ratio (Patterson and Moore 1986). Therefore, perceptual mechanisms that exploit spatial and temporal features of signals and noise could be particularly important in species that communicate in dense social aggregations of conspecifics. Accordingly, several studies of 'spatial unmasking' show that physical separation between target signals and background noise improves a receiver's ability to detect or recognize acoustic signals,

even when both have similar spectral content (reviewed in Bee and Micheyl 2008). In contrast, we still know little about perceptual mechanisms by which animals potentially exploit temporal features of signals and noise during signal recognition (Klump 1996; Hulse 2002; Bee and Micheyl 2008).

A well-known feature of natural sounds, including the noise generated by aggregations of signaling animals, is that their amplitudes fluctuate in time (Richards and Wiley 1980; Nelken et al. 1999; Vélez and Bee 2010). An ability to exploit level fluctuations in background noise contributes to human speech perception in noisy social settings. Speech recognition thresholds are usually lower when the level of speech-shaped noise (i.e., noise with the frequency spectrum of speech) fluctuates, compared to those when the same noise lacks level fluctuations (Gustafsson and Arlinger 1994; Bacon et al. 1998; Füllgrabe et al. 2006). Release from masking in temporally fluctuating noise is often referred to as 'dip listening' and is attributed to our ability to catch short 'acoustic glimpses' of target speech when the level of the background noise dips to low levels (Cooke 2006; Vestergaard et al. 2011). Masking release by means of dip listening is generally greater when noises fluctuate at slower, compared with faster, rates because the dips in noise levels are usually longer, increasing the probability of catching meaningful glimpses of target speech (Gustafsson and Arlinger 1994; Bacon et al. 1998). Importantly, however, level fluctuations in background noise do not always ameliorate speech recognition in noise. Psychophysical studies of 'modulation masking' in humans reveal that fluctuating maskers can impair recognition of temporally modulated target signals, particularly when the fluctuation rates of signal and noise are similar (Bacon and Grantham 1989). Because speech recognition heavily depends on the signal's temporal structure (Shannon et al. 1995), fluctuating background noise may sometimes impair speech recognition by means of modulation masking (Kwon and Turner 2001).

Psychophysical and neurophysiological studies have shown that level fluctuations in background noise can improve the ability of nonhuman animals to detect simple tonal signals and narrowband noises (Klump and Langemann 1995; Nelken et al. 1999; Langemann and Klump 2001, 2007; Nieder and Klump 2001; Hofer and Klump 2003; Bee et al. 2007; Jensen 2007; Branstetter and Finneran 2008; Fay 2011). However, only a few studies have investigated similar abilities in animals' recognition of their acoustic communication signals. Studies of Bow-winged grasshoppers (Chorthippus biguttulus; Ronacher and Hoffmann 2003) and Cope's gray treefrogs (Hyla chrysoscelis; Vélez and Bee 2011) indicate that receivers experience dip listening when recognizing temporally structured communication signals in slowly fluctuating noise backgrounds. In both species, signalers produce calls composed of a series of distinct pulses and pulse structure mediates species recognition (Helversen and Helversen 1997; Schul and Bush 2002). Importantly, both Bow-winged grasshoppers and Cope's gray treefrogs also experience modulation masking when masker fluctuation rates are similar to the pulse rates that mediate recognition of conspecific calls. These results raise important questions about the effects of fluctuating noise backgrounds in species with acoustic communication signals that do not comprise series of pulses.

Our primary objective in this study of green treefrogs (Hyla cinerea) was to investigate the extent to which dip listening and modulation masking affect signal recognition in a frog species with an advertisement call composed of a single note instead of a pulse train. Male green treefrogs form dense breeding choruses in which they produce a short (120–160 ms) advertisement call repeated about once or twice per second (Oldham and Gerhardt 1975). The advertisement call produced by males is composed of a harmonic series with spectral peaks around 900 and 3,000 Hz (Oldham and Gerhardt 1975). While the interaction of spectral components generates some amplitude modulations in the call, the call is delivered as a single note and not as a series of pulses. As in other anurans (Gerhardt and Huber 2002; reviewed in Vélez et al. in press), background chorus noise constrains signal recognition in green treefrogs (Gerhardt and Klump 1988a) and spatial unmasking leads to improvements in signal recognition by females (Schwartz and Gerhardt 1989).

We first recorded the sounds of green treefrog choruses and calculated their average modulation spectrum (Vélez and Bee 2010) to determine the rates at which chorus noise levels typically fluctuate. We next conducted phonotaxis experiments to test the overarching hypothesis that level fluctuations in background noise affect the ability of female frogs to recognize male advertisement calls. Together, our results suggest that fluctuation rates typical of the noise generated by green treefrog choruses should have little effect in terms of either facilitating (dip listening) or constraining (modulation masking) signal recognition by females in a chorus environment.

Materials and methods

Study sites and subjects

We made acoustic recordings and collected females in amplexus from artificial ponds at the Jasper State Fish Hatchery in Jasper, Texas (30°57′ 4.01″ N, 94°7′ 39.53″ W), between 2200 and 0200 h in April and May 2010 and 2011, and July 2010. Females were kept at approximately 2–4 °C

in ice-filled coolers to delay oviposition and were transported overnight (within 24–48 h of collection) to the St. Paul campus of the University of Minnesota, where they were tested using the same general protocols and under the same general conditions as in our previous study of dip listening and modulation masking in Cope's gray treefrogs (Vélez and Bee 2011). At least 1 h prior to testing, females were placed in a 24 °C incubator to allow their body temperature to reach 24 °C ± 1 °C. At the completion of testing, subjects were returned to the pond from which they were collected.

Acoustical analysis: level fluctuations of natural chorus noise

Digital recordings (44,100 Hz sampling rate, 16-bit resolution) were made using a Marantz PMD 670 recorder and an omnidirectional Sennheiser ME62 microphone. We positioned the microphone 5 cm above ground level at distances ranging between 5 and 15 m from the nearest calling male. We chose a position close to the ground because females at our field site usually approach choruses from such positions. The range of distances to the nearest calling male was chosen to minimize the predominance of calls with high signal-to-noise ratios produced by a single nearby male. Although we did not attempt to quantify the densities of calling males, we made recordings during peaks of calling activity. We chose for analysis 25 90-s segments of green treefrog choruses, each one taken from chorus recordings made on different days at one of four different ponds, or at different times and locations within a pond (3–11 recordings per pond). Air, water, and wet-bulb air temperatures at the times of recording ranged between 19.8 and 26.2 °C ($\bar{X}$ ± SD = 24.3 ± 1.9 °C), 22.6 and 32 °C ($\bar{X}$ ± SD = 27.1 ± 3.2 °C), and 18.8 and 26.8 °C ($\bar{X}$ ± SD = 23.9 ± 2.5 °C), respectively. We only used recordings in which other species were absent or calling at very low densities and away from the microphone (overall amplitude of heterospecific signals relative to green treefrog chorus noise less than −6 dB).

Chorus recordings were analyzed using custom-written scripts in Matlab v7.6 (Mathworks, MA, USA). We estimated level fluctuation rates by calculating a modulation spectrum for each chorus recording, which represents the power spectrum of the envelope of the chorus noise. To this end, we first down-sampled the 90-s recorded segments to a 11,025 Hz sampling rate and normalized them to the same root-mean-square (RMS) amplitude to compensate for differences in microphone positions and recorder gain settings at the times recordings were made. We then extracted the Hilbert envelope of the waveform and, to correct for the DC offset, subtracted the mean value of the

envelope from each sample of the envelope. To determine the rates of level fluctuation, we calculated the long-term modulation spectrum of each 90-s segment as the fast-Fourier transform (FFT) of its envelope (sampling rate = 11,025 samples/s, Hamming window size = 65,536 points, overlap = 25 %). We normalized each spectrum to the maximum value of the magnitude of its FFT and then calculated the mean and standard deviation of the 25 modulation spectra. Mean and ±1 standard deviation modulation spectra were transformed to a dB scale [$20*\log_{10}$(FFT magnitude)] and smoothed for plotting purposes using a running average of 11 points.

General testing procedures

The phonotaxis behavior of green treefrogs has been well described in previous studies (reviewed in Gerhardt 2001). Our general testing procedures also have been described previously and readers are referred to those studies for additional details (Bee and Schwartz 2009; Vélez and Bee 2011). Briefly, we conducted single-speaker phonotaxis experiments (Gerhardt 1995) under infrared (IR) illumination in a circular test arena (2 m diameter) made of acoustically transparent but visually opaque walls (60 cm height). The test arena was placed inside a walk-in, temperature-controlled (24 °C ± 1 °C), hemi-anechoic sound chamber (details in Bee and Schwartz 2009). Using an IR-sensitive video camera mounted above the center of the test arena, behavioral responses were simultaneously encoded to digital video files and scored in real time by two observers on a monitor outside the chamber.

We divided the perimeter of the test arena into 24 15° arcs. The speaker used to broadcast target signals (a/d/s/ L210 or Orb Mod1) was placed on the floor just outside the wall of the arena, 1 m away from a release point at the center of the arena, and centered in one of the 15° arcs. To eliminate any possibility of directional response bias, the position of the target speaker was randomly varied around the arena's perimeter between tests of two to four females. Maskers were broadcast from a Kenwood KFC-1680ie speaker suspended from the ceiling of the chamber 190 cm above the central release point. This overhead speaker created a fairly uniform (±2 dB) sound field across the floor of the circular test arena. We used a Brüel and Kjær Type 2250 sound level meter to measure and calibrate sound levels at the approximate position of a subject's head at the central release point. Our playback system had a flat (±3 dB) frequency response between 500 and 5,000 Hz when measured at the central release site using the Brüel and Kjær sound level meter and a Stanford Research SR780 spectrum analyzer.

At the beginning of each test, we placed a subject in an acoustically transparent holding cage located at the arena's

central release point. After a 1.5-min silent acclimation period, broadcasts of the target signal were initiated. Subjects were released after 15 s of signal presentation using a rope and pulley system operated from outside the chamber. In tests in which a masking noise was used, broadcasts of the masker began 30 s before the onset of the target signal and continued throughout the entire test. Each subject was used in a series of 6–17 phonotaxis tests and was given a timeout period of 5–15 min inside the incubator between tests. Unless noted otherwise, we scored a correct response in a phonotaxis test if (i) the subject first touched the wall of the test arena in the hemi-circle containing the speaker broadcasting the target signal, (ii) the subject made contact with the arena's wall inside the 15° bin in front of the speaker broadcasting the target signal within 5 min of being released, and (iii) the subject remained for 30 s in a bin of approximately 30° width centered in front of the speaker.

Acoustic stimuli

All acoustic stimuli used in this study were generated in Matlab v7.6 (Mathworks, MA, USA) with a sampling rate of 11,025 Hz and 16-bit resolution.

The standard call

The target signal was a standard synthetic call with values of temporal and spectral properties close to the averages of green treefrog calls recorded at our study site (Höbel unpublished data; Fig. 1a, b). The standard call consisted of three phase-locked sinusoids with equal relative amplitudes and frequencies of 850, 2,550, and 2,833 Hz. The call was 147 ms in duration and repeated at a rate of 1.78 calls/s (562 ms call period). The amplitude-envelope of the call was shaped with a 25-ms rise time and a 50-ms fall time (50 % amplitudes were reached at 25 % of the rise time and at 75 % of the fall time).

Chorus-shaped maskers

We created chorus-shaped maskers by filtering white noise so that the long-term frequency spectrum resembled that of natural green treefrog choruses (Fig. 1a). As models of chorus noise, we selected 10 100-s recordings of green treefrog choruses obtained at our study site near the peak of nightly vocal activity using the recording protocols described above. Recordings were down-sampled to 11,025 Hz, normalized to a common RMS amplitude, and transformed to the spectral domain (FFT size = 4,194,304). We created
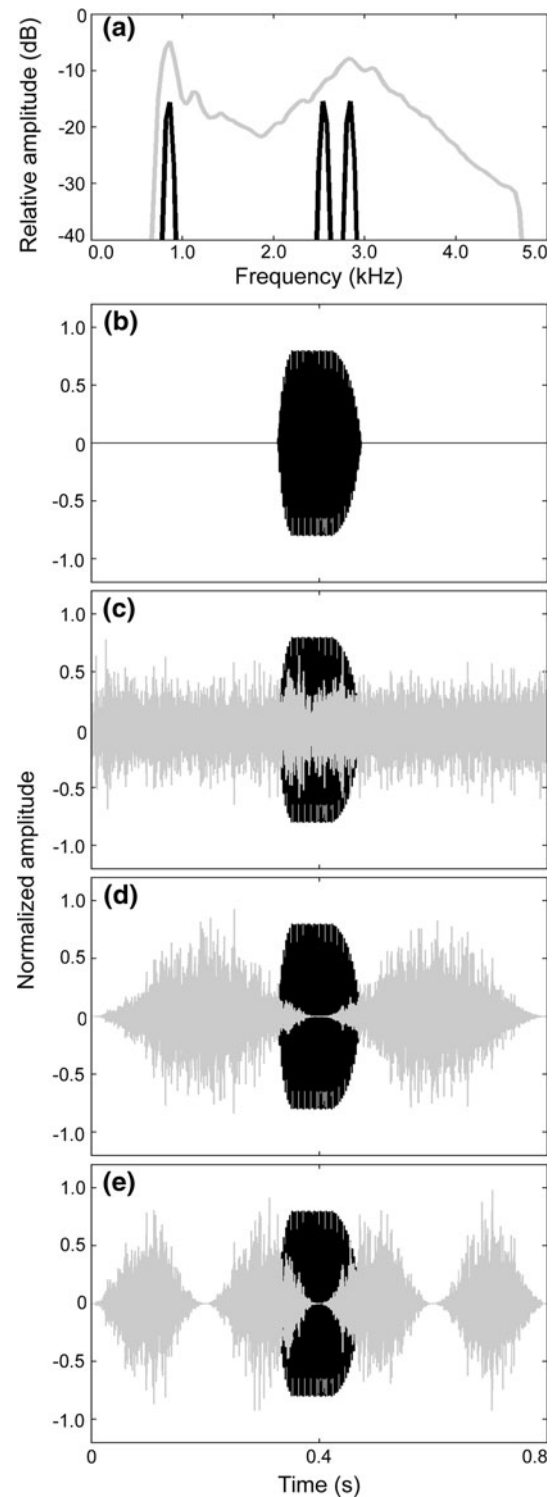

Fig. 1 **a** Frequency spectra of the standard call used as a target signal (*black line*) and the chorus-shaped maskers (*gray line*) calculated with 256-point Hamming windows. Waveforms of 0.8 s segments of the standard call alone (**b**), and in the presence of a non-fluctuating masker (**c**), a 2.5 Hz SAM masker (**d**), and a 5 Hz SAM masker

an exemplar of chorus-shaped noise by generating a 6-min white noise (sampling rate = 11,025 Hz), transforming it to the spectral domain (FFT size = 4,194,304), and multiplying its frequency spectrum by the average frequency spectrum determined from our 10 recordings of natural choruses. The resulting frequency spectrum was band-pass filtered between 700 and 4,700 Hz and transformed to the time domain with an inverse FFT. The resulting noise had a frequency spectrum typical of green treefrog choruses and a flat temporal envelope with no amplitude fluctuations beyond the random inherent fluctuations typical of band-limited noise (Fig. 1c). We refer to this type of noise as 'non-fluctuating' chorus-shaped noise because we did not impose any modulations on the resulting waveforms. For our experiments, we created four exemplars of non-fluctuating chorus-shaped noises using a different white noise for each.

Fluctuating maskers were generated by imposing sinusoidal amplitude modulation (SAM) at rates of 0.625, 1.25, 2.5 (Fig. 1d), 5 (Fig. 1e), 10, 20, 40, and 80 Hz on all four exemplars of non-fluctuating chorus-shaped noise. The use of SAM noise is common in psychophysical and neurophysiological studies of temporal auditory processing due to the high level of experimental control they offer (Joris et al. 2004). We created SAM noises by multiplying the non-fluctuating chorus-shaped noises by a modulating sine wave following the equation:

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

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 (from 0.625 to 80 Hz in one-octave steps), ρ is the starting phase of the modulator, $c(t)$ is the non-fluctuating chorus-shaped noise, and t is time in s. The starting phase of the modulator differed for each of the four exemplars ($\rho = 0°$, $90°$, $180°$, or $270°$). We used a different starting phase for each exemplar because, during phonotaxis tests, the number of calls falling within dips of fluctuating maskers could depend on the starting phase of the masker and the timing of the target signal, potentially affecting signal recognition. Equal numbers of subjects were tested with each exemplar.

Experiment 1: call recognition in fluctuating chorus-shaped maskers

In two previous studies of green treefrogs, call recognition thresholds were estimated in the presence (Gerhardt and Klump 1988a) and absence (Gerhardt 1981) of chorus noise using two different procedures. Therefore, our first objective in this experiment was to measure signal recognition thresholds in the presence and absence of non-fluctuating chorus-shaped maskers using the same standardized protocol (Bee and Schwartz 2009).

The second, and main, objective of this experiment was to determine the extent to which level fluctuations in background noise affect the ability of female green treefrogs to recognize advertisement calls. We tested the null hypothesis that level fluctuations do not affect signal recognition thresholds against two alternative hypotheses: the dip-listening hypothesis and the modulation-masking hypothesis. According to the dip-listening hypothesis, we predicted lower signal recognition thresholds in the presence of fluctuating maskers compared with those measured in the presence of non-fluctuating maskers. In contrast, if females experienced modulation masking, we predicted higher signal recognition thresholds in the presence of fluctuating maskers compared with those measured using non-fluctuating maskers.

Experimental design

Using a between-subjects experimental design, we tested 24 subjects in each of ten randomly assigned treatments (total $N = 240$). A 'no-masker' treatment, in which no masking noise was broadcast, served as a control to measure signal recognition thresholds in the absence of background noise. This control treatment allowed us to evaluate the effects of our chorus-shaped maskers on subjects' responses to the target signal. (We describe generation of signal recognition thresholds in the next section). In a second treatment, we determined thresholds in the presence of non-fluctuating chorus-shaped noise broadcast from the overhead speaker. This 'non-fluctuating noise' treatment served two purposes. First, we used this treatment to estimate the difference in signal recognition thresholds in the presence and absence of chorus-shaped noise. Second, this treatment served as a control to assess the effects of level fluctuations in masking noise on subjects' ability to recognize the target signal. In the remaining eight treatments, we measured thresholds in the presence of one of the eight SAM chorus-shaped maskers (i.e., 0.625–80 Hz SAM in octave steps) broadcast form the overhead speaker. We refer to these eight treatments as 'fluctuating noise' treatments. In all nine treatments that involved broadcasts of a chorus-shaped masker, the equivalent long-term RMS amplitude of the masker was calibrated at the central release point of the arena to a sound pressure level (SPL re. 20 μPa) of 73 dB (LC_{eq}). This level falls within the range of chorus noise levels measured in the field (Vélez and Bee unpublished data).

Signal recognition thresholds

As in previous studies (Bee and Schwartz 2009; Vélez and Bee 2011; Nityananda and Bee 2012), we operationally defined the signal recognition threshold as the lowest signal

level necessary to elicit positive phonotaxis to the target signal. As defined here, signal recognition thresholds differ from traditional signal detection thresholds because positive phonotaxis requires that subjects (i) detect the signal, (ii) recognize it as a conspecific advertisement call, and (iii) localize its source. Following Bee and Schwartz (2009), we estimated signal recognition thresholds using an adaptive tracking procedure in which subjects were tested in a series of 'reference,' 'sham,' and 'test' trials. The total number of trials in a series depended on the subjects' behavioral responses and ranged between six and 17 across all 240 subjects. Each series of trials began and ended with a reference trial. In reference trials, the target signal was the standard call broadcast at 85 dB SPL (LCF) in the absence of masking noise. This signal level corresponds to natural call amplitudes measured at 1 m (Gerhardt 1975). Subjects were also tested in a reference trial if they failed to meet our response criteria in any two consecutive test trials. Subjects failing to meet the response criteria in any reference trial were replaced and their data were discarded. This procedure ensures that all subjects included in the analyses were motivated to respond throughout all trials and is necessary to validate test trials in which subjects fail to respond to the target signal (Bush et al. 2002).

The trial following the first reference trial was a sham trial, in which we tested the subject without presenting a target signal (Sham trials are described in more detail in a subsequent section). Following the sham trial, subjects began a series of test trials. In the first test trial of a series, the target signal was broadcast at 48 dB SPL (LCF) in the no-masker treatment and 73 dB SPL (LCF) in all treatments with a masker. These levels were selected based on previous estimates of signal recognition thresholds in the presence and absence of noise for our study species (Gerhardt 1981; Gerhardt and Klump 1988a). In all subsequent test trials, the level of the target signal was systematically varied contingent upon the subject's behavior in the previous test trial. If the subject responded in a test trial, the level of the signal was decreased by 3 dB on the next test trial. If, on the other hand, the subject failed to respond in a test trial, the level of the target signal was increased by 3 dB on the next test trial. We continued this pattern until there was a change in the subject's behavior between two consecutive test trials (i.e. from response to no response, or vice versa). After a subject changed its behavior between two test trials, we conducted a final test trial in which the direction of signal level change was reversed and the step size was reduced to 1.5 dB. We calculated a subject's recognition threshold as the average between the minimum signal level that elicited positive phonotaxis and the maximum signal level that failed to do so (Bee and Schwartz 2009). Signal recognition thresholds were calculated in a linear scale and converted back to decibels.

Statistical analyses

We compared signal recognition thresholds in the presence and absence of non-fluctuating chorus-shaped maskers with a Mann–Whitney U test. Signal recognition thresholds in these two treatments were normally distributed (Kolmogorov–Smirnov tests $Ps > 0.2$) but did not meet the homogeneity of variance assumption for parametric testing (Levene's test: $F_{1,46} = 18.827$, $P < 0.001$).

To examine the extent to which level fluctuations affected masked signal recognition thresholds, we first analyzed thresholds in the non-fluctuating control treatment and all the fluctuating noise treatments using a one-way analysis of variance (ANOVA). Signal recognition thresholds in these treatments were normally distributed (Kolmogorov–Smirnov tests $Ps > 0.05$) and homoscedastic (Levene's test: $F_{8,207} = 0.78$, $P = 0.62$). We treated starting phase of each exemplar as a between-subjects factor in preliminary statistical analyses. Neither 'starting phase' nor the interaction between 'starting phase' and 'masking treatment' had a significant effect on signal recognition thresholds; therefore, we dropped these factors from the ANOVA models reported below (starting phase: $F_{3,180} = 1.36$, $P = 0.26$, partial $\eta^2 = 0.02$; starting phase × masking treatment: $F_{24,180} = 0.44$, $P = 0.99$, partial $\eta^2 = 0.06$). We used planned contrasts to test the dip-listening and modulation-masking hypotheses. Compared to thresholds measured in non-fluctuating maskers, the dip listening hypothesis predicts relatively *lower* thresholds in the presence of fluctuating maskers, whereas the modulation-masking hypothesis predicts relatively *higher* thresholds in fluctuating maskers. A significance criterion of $\alpha = 0.05$ was used in all tests.

Sham trials: subject behavior in the absence of target signals

We tested females in sham trials to observe their behavior in our experimental apparatus in the presence of noise but in the absence of a specific target signal. This control treatment was important because the 'noise' generated in a chorus can actually serve as a 'signal' that allows individuals of some species of frogs to locate breeding sites (Gerhardt and Klump 1988b; Bee 2007; Swanson et al. 2007; Christie et al. 2010). Consequently, chorus sounds and chorus-like noises potentially affect subject behavior during signal recognition experiments. In the present experiment, such behavior could have biased our results if subjects behaved in ways that affected estimates of recognition thresholds (e.g., by exhibiting more active searching behavior) in the presence of some, but not all, of the maskers. Sham trials allowed us to evaluate this possibility.

Note that sham trials in the no-masker treatment were designed to provide a baseline of subject behavior in the sound chamber without any acoustic stimulation. During sham trials of all other treatments, each masker was broadcast continuously from the overhead speaker with a long-term RMS amplitude of 73 dB SPL (LC_{eq}). Since no stimuli were broadcast from speakers on the floor, there was no 'correct' response in any sham trials. Therefore, we did not apply any response criteria in sham trials and recorded female behavior for 5 min after her release.

We used two approaches to test the null hypothesis that movement patterns were not affected by the presence/absence or type of masker broadcast from the overhead speaker. First, we compared the proportion of subjects that left the holding cage and showed any type of movement in the arena. Our reasoning was that, if the sounds of some chorus-shaped maskers caused the subjects to show either more or fewer overall movements, we would find a difference in the proportion of females leaving the holding cage across treatments. We used Fisher's exact tests to test the null hypothesis that the proportion of females leaving the holding cage in each masking treatment was not different from the proportion of females doing so in the no-masker treatment (i.e., when there was no sound present at all). Second, we used the animal tracking software EthoVision v3.1 (Noldus 2005) to compare movement patterns of the subjects that left the holding cage. We measured the total distance moved by the subjects (in cm), the average velocity of their movements (in cm/s), the average turn angles (in degrees) and the meander (in degrees/cm), which quantifies the magnitude of changes in direction of movements relative to the distance moved (see Bee and Riemersma 2008). Because these four variables were intercorrelated, we used principal component analysis (PCA) to examine the extent of covariation and to obtain independent factors describing movement patterns. We investigated movement patterns across treatments using Kruskal–Wallis tests.

Experiment 2: chorus-shaped maskers as potential signals

As previously mentioned, natural chorus sounds may act as signals used by individuals to locate breeding sites (Gerhardt and Klump 1988b; Bee 2007; Swanson et al. 2007; Christie et al. 2010). We have also found this to be true for artificially generated noises that have either the natural frequency spectrum of a chorus and unnatural (e.g., sinusoidal) envelope fluctuations or an unnatural spectrum (e.g., broadband noise) with natural envelope fluctuations (see below; Vélez and Bee 2012). Ambient sounds can also be used by some frogs to evaluate potential threats. For instance, the temporal structure of the cracking sounds of

burning vegetation appears to be an important cue used by juvenile reed frogs (*Hyperolius nitidulus*) to flee from approaching fire (Grafe et al. 2002). Consequently, temporally structured sounds other than individual communication signals can have behavioral salience to frogs and potentially 'compete' with the target signal in signal recognition experiments. We evaluated this possibility in a control experiment in which we tested the null hypothesis that the noises used as maskers in experiment 1 fail to elicit either positive or negative phonotaxis when broadcast as potential target signals.

Using a within-subject experimental design, we tested 20 subjects in nine treatments, each corresponding to one of the chorus-shaped noises used in experiment 1. Individual females were tested in a series of 11 trials: two reference trials (initial and final) and nine intervening test trials. In each test trial, a chorus-shaped noise was broadcast continuously from a speaker on the floor just outside the arena wall. We calibrated the long-term RMS amplitude of the chorus-shaped noises to 73 dB SPL (LCeq) at the central release point of the test arena. The 15° bin where the speaker was placed was designated 0°, and we measured the angle at which females first touched the wall of the arena in 15° bins. The position of the target speaker was randomly varied between groups of two or three subjects. As in previous similar studies (Swanson et al. 2007; Vélez and Bee 2010, 2011), we ended the test trials as soon as subjects made contact with the wall anywhere around the test arena or after 5 min of stimulus presentation had elapsed. For each subject, we randomized the order in which the chorus-shaped noises were presented. We tested the null hypothesis that orientation angles were uniformly distributed using Rayleigh tests. We reasoned that, if subjects were attracted to or repelled by the chorus-shaped noises, orientation angles would not be uniformly distributed, but would be directed toward or away from the stimulus, respectively.

Results

Acoustical analysis: level fluctuations of natural chorus noise

The noise generated by green treefrog choruses fluctuates in level (Fig. 2a, b, c). The average modulation spectrum of green treefrog choruses is characterized by low-frequency modulation rates (below approximately 10 Hz) with a prominent peak around 5 Hz (Fig. 2d). Analyses of modulation spectra also revealed a second peak of level fluctuations around 300 Hz, with an amplitude of approximately −18 dB relative to the dominant peak at 5 Hz (Fig. 2d).
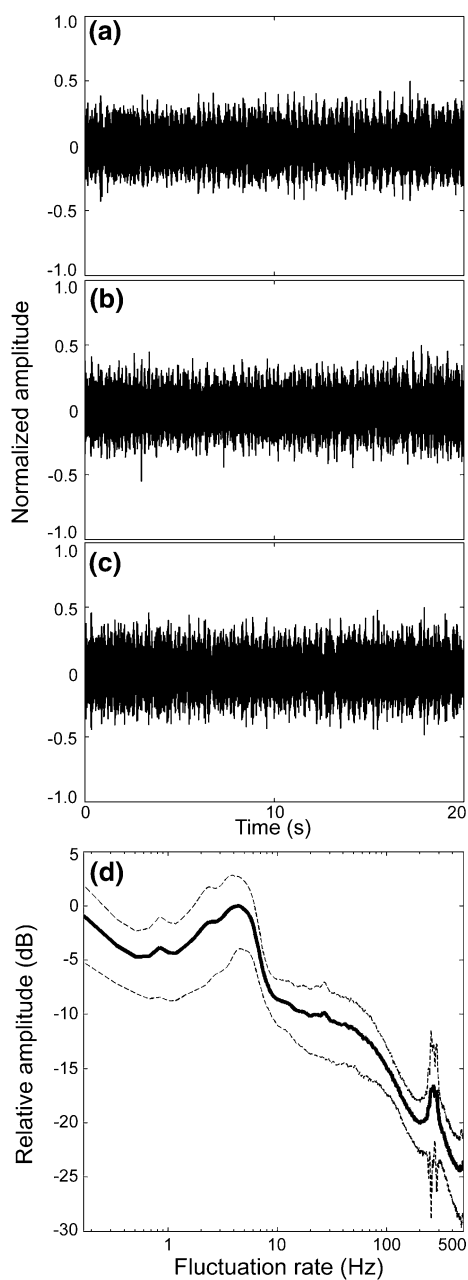
Fig. 2 **a**, **b**, **c** Waveforms of three representative green treefrog chorus recordings. **d** Mean (*solid line*) ±1 standard deviation (*dashed lines*) modulation spectrum calculated from 25 90-s segments of green treefrog chorus recordings shown on a logarithmic *x* axis
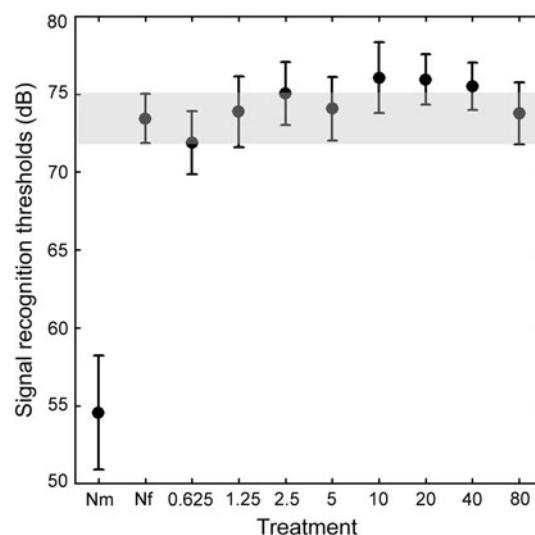

Fig. 3 Mean (±95 % confidence intervals) signal recognition thresholds as a function of masking treatment. The *gray box* serves to illustrate the ±95 % confidence intervals in the non-fluctuating control treatment. *Nm* no-masker treatment, *Nf* non-fluctuating treatment. 0.625–80 represent SAM rates (in Hz) of each fluctuating noise treatment

Experiment 1: call recognition in fluctuating chorus-shaped maskers

Signal recognition thresholds

Signal recognition thresholds were significantly higher (by ∼20 dB) in the presence of the non-fluctuating masker than in the absence of noise (Man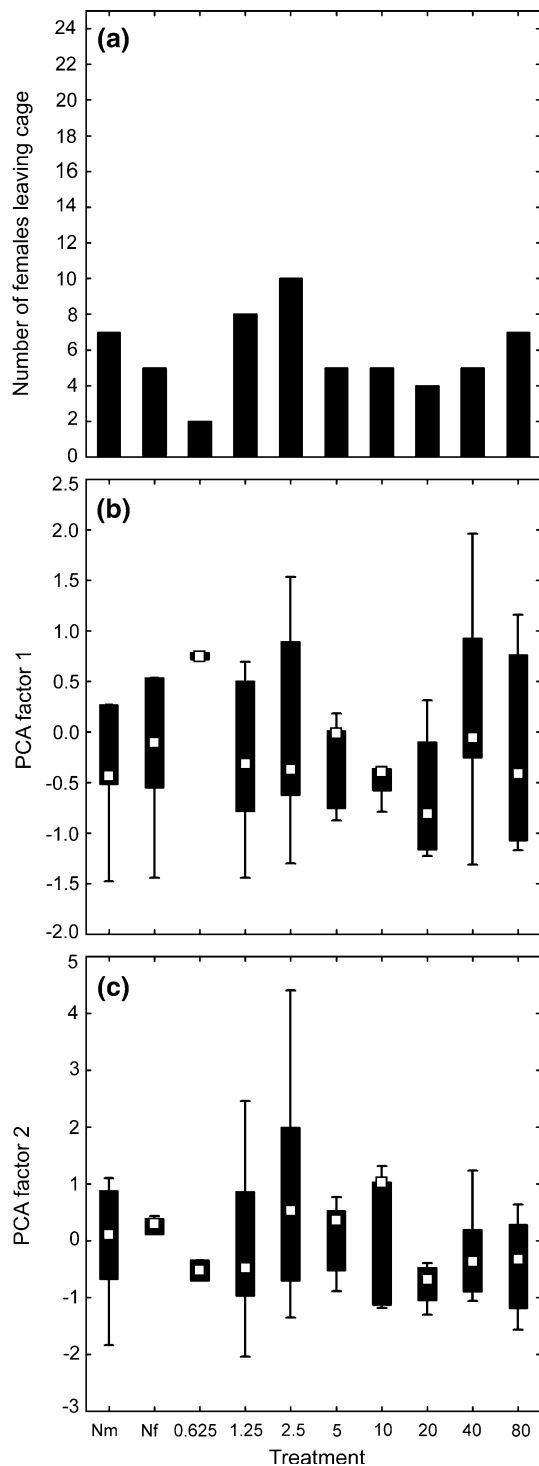n–Whitney *U* Test: $U = 19.0$, $P < 0.0001$). The median signal recognition threshold was 53.3 dB (mean = 54.6 dB; Fig. 3) in the no-masker control treatment and 73.8 dB (mean = 73.4; Fig. 3) dB in the non-fluctuating control treatment.

On average, signal recognition thresholds in the presence of fluctuating maskers ranged from about 1.5 dB lower to 2.5 dB higher than the average recognition threshold in the non-fluctuating control treatment (Fig. 3). We found a significant effect of masker type on signal recognition thresholds (one-way ANOVA: $F_{8,207} = 2.09$, $P = 0.038$, $\eta^2 = 0.08$). Individual planned comparisons failed to reveal any significant differences between signal recognition thresholds in the non-fluctuating control treatment and each fluctuating noise treatment; these differences approached statistical significance at masker fluctuation rates of 10 Hz ($F_{1,207} = 3.87$, $P = 0.0504$, partial $\eta^2 = 0.018$) and 20 Hz ($F_{1,207} = 3.51$, $P = 0.0623$, partial $\eta^2 = 0.017$), but not at any of the other fluctuation rates (all $Fs_{1,207} < 2.4$, $Ps > 0.1$, partial $\eta^2 s < 0.012$).

Sham trials: subject behavior in the absence of target signals

We found no significant differences in movement patterns between masking treatments in the sham trials (Fig. 4). Seven of 24 subjects (29.2 %) left the holding cage during the sham trial in the no-masker treatment. In treatments with masking noise, between two (8.3 %) and 10 (41.7 %) of 24 subjects left the holding cage during the sham trials (Fig. 4). The proportions of females that left the holding cage in treatments that included a masker were not

PCA factor accounted for 49.04 % of the total variance, had an eigenvalue of 1.96, and loaded most heavily on variables related to angular movements: meander (factor-variable correlation = 0.96), mean turn angle (0.95), mean velocity (−0.31), and total distance moved (0.18). The second PCA factor had an eigenvalue of 1.37, described an additional 34.16 % of the total variance, and loaded most heavily on variables related to non-angular movements: total distance moved (0.84), mean velocity (0.79), mean turn angle (0.16), and meander (−0.07). We found no differences across treatments in PCA factor 1 (angular movements; Kruskal–Wallis test: $H_{9,58} = 5.28$, $P = 0.81$) or PCA factor 2 (non-angular movements; $H_{9,58} = 5.82$, $P = 0.76$) during sham trials (Fig. 4). Parallel analyses on the original movement-pattern variables confirmed these results (Supplementary material 1).

Experiment 2: chorus-shaped maskers as potential signals

Results from this control experiment are summarized in Table 1. Subjects were strongly oriented to the target speaker in reference trials, as evidenced by mean vector angles close to 0° and lengths of mean vectors close to 1.0 (Rayleigh test: $Zs > 18$, $Ps < 0.001$; Table 1). Subjects were neither attracted to, nor repelled by, the non-fluctuating noise or by noises that fluctuated at rates of 0.625, 10, 20, 40, and 80 Hz, as evidenced by uniformly distributed orientation angles ($Ps \geq 0.09$; Table 1). Orientation angles, however, were not uniformly distributed when chorus-shaped noises fluctuated at rates of 1.25, 2.5, and 5 Hz ($Ps < 0.01$; Table 1). Mean angles close to 0° and lengths of mean vectors greater than 0.6 suggest that the females that touched the wall of the arena were attracted to these chorus-shaped noises.

Discussion

In this study of green treefrogs, we found an increase of approximately 20 dB in signal recognition thresholds in the

significantly different from that in the no-masker treatment (two-tailed Fisher's exact tests: all $Ps > 0.13$).

The four variables of movement patterns measured (i.e., total distance moved, average velocity, average turn angle, meander) could be summarized by two principal components with eigenvalues greater than 1.0, which together explained 83.2 % of the total variance. The first

Table 1 Results of circular analyses for orientation angles at the arena wall in response to reference trials and sinusoidally amplitude-modulated (SAM) chorus-shaped noises

Treatment	Circular analysis					
	N	Mean vector (°)	Length of mean vector (r)	Circular standard deviation (°)	Rayleigh test Z	Rayleigh test P
Reference 1	20	0	0.99	5	19.9	<0.01
Non-fluctuating	1	135	1.00	–	1.0	0.51
0.625 Hz SAM	3	249	0.70	48	1.5	0.25
1.25 Hz SAM	13	3	0.62	56	4.9	<0.01
2.5 Hz SAM	14	3	0.96	16	12.9	<0.01
5 Hz SAM	11	7	0.81	37	7.3	<0.01
10 Hz SAM	3	341	0.88	29	2.3	0.09
20 Hz SAM	5	226	0.21	101	0.2	0.81
40 Hz SAM	5	220	0.64	54	2.0	0.13
80 Hz SAM	3	115	0.45	73	0.6	0.59
Reference 2	20	4	0.96	16	18.5	<0.01

N represents the number of subjects that touched the wall of the arena in each treatment; the total number of subjects tested per treatment was 20

presence of chorus-shaped maskers broadcast at natural levels. We also found that level fluctuations in background noise had little effect on the ability of females to recognize male advertisement calls. Mean signal recognition thresholds in fluctuating noise treatments were never statistically different from those in the non-fluctuating control treatment. These results are thus inconsistent with both the dip-listening and the modulation-masking hypotheses. Studies that report null results, like this one, are often open to criticisms regarding inadequate sample sizes (and thus, low statistical power), inappropriate experimental designs, or both. Therefore, we consider these issues next.

Statistical power and experimental design

Statistical power refers to the probability of rejecting a false null hypothesis and varies as a function of effect size and sample size at a given alpha level (Cohen 1988; Rosenthal and Rosnow 1991). A common effect size for the F statistic from ANOVA is partial η^2, which describes the proportion of total variance in the dependent variable accounted for by a given factor. Our study had adequate statistical power to detect a relatively small overall effect of masker type on signal recognition thresholds (partial $\eta^2 = 0.08$). Even with our relatively large sample size (total $N = 240$), however, the statistical power of our focused contrasts was not high enough to detect the very small effects corresponding to differences between the non-fluctuating control treatment and each of the fluctuating noise treatments. The differences between mean signal recognition thresholds in the non-fluctuating control treatment and all fluctuating noise treatments ranged between 0.3 and 2.6 dB, and the corresponding statistical effect

sizes (partial η^2) ranged from less than 0.001 to 0.018 (Supplementary material 2). These effect sizes are considerably smaller than some of those found in a previous study of Cope's gray treefrogs using the same experimental design, sample size, and statistical analyses (Vélez and Bee 2011; Supplementary material 2). In Cope's gray treefrogs, mean threshold differences between the non-fluctuating control treatment and all fluctuating noise treatments ranged from 0.4 to 5.6 dB. The corresponding statistical effect sizes ranged between 0.001 and 0.096. Therefore, the sample size of the present study provided sufficient statistical power to detect the small to moderate effect sizes observed in Cope's gray treefrogs, but not the smaller effects reported here for green treefrogs. Together, our results suggest the effects of temporally fluctuating maskers on signal recognition thresholds are smaller in green treefrogs compared with Cope's gray treefrogs. Previous studies suggest that differences of 2–4 dB in signal levels are biologically important for females listening to potential mates in natural settings (Gerhardt et al. 2000; Vélez and Bee 2011; Bee et al. 2012). Smaller effects, as the ones we report for green treefrogs, are less likely to be biologically meaningful in the din of a breeding chorus.

Regarding the experimental design, we followed the same protocol as in Vélez and Bee (2011), which proved adequate to demonstrate both dip listening and modulation masking in Cope's gray treefrogs. There is, however, one caveat potentially affecting interpretations of results from the present study of green treefrogs: three of the eight fluctuating maskers (1.25, 2.5 and 5 Hz) were behaviorally salient when presented as potential target signals in experiment 2 (Table 1). Consequently, these sounds had some potential to 'compete' with the target signals used to

measure signal recognition thresholds in experiment 1. Interpreting results from the 1.25, 2.5 and 5 Hz treatments of experiment 1, therefore, should be done with some caution. For two reasons, we believe the degree of 'competition' for subjects' responses between these three maskers and the standard call in experiment 1 was likely small. First, across the 1.25, 2.5 and 5 Hz treatments of experiment 2, only between 55 and 70 % of subjects reached the arena wall within 5 min (Table 1). By comparison, approximately 96 % of females did so in response to the standard call presented at an equivalent (i.e., 73 dB) or *lower* level in the no-masker treatment of experiment 1 (data not shown). Hence, the 1.25, 2.5 and 5 Hz maskers should be regarded as marginally attractive at best. Second, and perhaps more importantly, results from the sham trials of experiment 1 showed that subject movement patterns were not differentially affected by the type of chorus-shaped noise broadcast from the overhead speaker (Fig. 4, Supplemental material 1). Together, these results suggest that the potential confounding effects of the 1.25, 2.5 and 5 Hz maskers used in measuring signal recognition thresholds in experiment 1 were probably small. Nevertheless, we cannot completely rule out the possibility that these three maskers competed with the target signal, affecting our estimates of signal recognition thresholds in these three (of eight) treatments. With this caveat in mind, it is worth discussing how our results relate to those of previous studies and the implications of our results for communication in natural environments in green treefrogs.

Comparisons with previous studies

Signal recognition in green treefrogs

In the absence of noise, we estimated an average signal recognition threshold of 54.6 dB (median = 53.3 dB) using an adaptive tracking method. Gerhardt (1981) tested females at different fixed signal levels and reported a behavioral recognition threshold of 48 dB. The difference between the two threshold estimates is likely due to methodological differences. Gerhardt (1981) determined 48 dB as a reasonable behavioral threshold because approximately 36 % of the females approached the target speaker within 30 cm at this signal level. In our experiment, 33 % of the females tested approached the speaker within 30 cm at a signal level of 48 dB. Therefore, even though the two criteria yield different threshold estimates, female behavior in both studies was nevertheless quite similar. Our estimates of signal recognition thresholds in the presence of background noise, when expressed as signal-to-noise ratios, are also similar to those reported previously. In the presence of non-fluctuating chorus-shaped maskers broadcast at 73 dB SPL (LC_{eq}), we estimated an

average signal recognition threshold of 73.4 dB. This masked signal recognition threshold corresponds to a signal-to-noise ratio of 0.4 dB. Using a short recording of chorus noise as a masker, Gerhardt and Klump (1988a) determined that female green treefrogs recognize calls embedded in background noise when the signal-to-noise ratio was at least 0 dB. While there is a difference in the temporal structure between our non-fluctuating chorus-shaped masker and their chorus recording, it may not have a strong effect on signal recognition, as will be discussed shortly. Our results corroborate Gerhardt and Klump's (1988a) conclusion that males that produce calls at sound levels lower than the level of the ambient chorus noise have little chance of attracting females.

Dip listening and modulation masking in nonhuman communication systems

Dip listening and modulation masking in the context of nonhuman animal communication have only been investigated in Bow-winged grasshoppers (Ronacher and Hoffmann 2003) and Cope's gray tree frogs (Vélez and Bee 2010, 2011). Female Bow-winged grasshoppers produce stridulatory mating signals composed of a series of pulsed syllables. In the presence of slowly fluctuating SAM maskers, male Bow-winged grasshoppers experienced masking release when the target signal was long (12 syllables), but not when it was short (4 syllables). These results offer partial support for the dip-listening hypothesis and suggest that signal duration affects signal recognition in fluctuating noise in grasshoppers. Interestingly, males experienced 5–10 dB of modulation masking with both short and long signals when background noise levels fluctuated at rates similar to those in the mating signal.

Likewise, females of Cope's gray treefrogs exhibit similar patterns of dip listening and modulation masking. Male gray treefrogs produce pulsed advertisement calls (40–60 Hz pulse rate) repeated about once every 5 s. In a first study with Cope's gray treefrog females, Vélez and Bee (2010) found an effect of modulation masking of approximately 12 dB in the presence of 40 Hz SAM maskers. In a subsequent study, female Cope's gray treefrogs experienced a release from masking of 2 to 4 dB in the presence of SAM chorus-shaped maskers that fluctuated at rates below 5 Hz, and about 4 dB to 6 dB of modulation masking when maskers fluctuated at rates above 20 Hz (Vélez and Bee 2011). Masking release in gray treefrogs in slowly fluctuating backgrounds was attributed to the ability of female frogs to catch meaningful acoustic glimpses of the advertisement call. Nine or more consecutive pulses of the gray treefrog call fell within dips (measured at the 6-dB downpoints) of maskers fluctuating at rates of 2.5 Hz or slower. Vélez and Bee (2011) showed

that, in the absence of background noise, calls with nine or more consecutive pulses were necessary to elicit phonotaxis. The strong effect of modulation masking observed in gray treefrogs in the 40 and 80 Hz SAM treatments was attributed to the interference caused by background noise level fluctuations at rates similar to those in the pulsed call. This result is consistent with other studies of call interference in gray treefrogs showing that overlapping calls disrupt perception of pulse rate by females (Marshall et al. 2006; Schwartz and Marshall 2006).

In contrast, our results with green treefrogs offer little support for the operation of dip listening or modulation masking in this species. In the present study, the maximum number of consecutive calls (call duration = 147 ms, call period = 562 ms) falling between the 6-dB down points of dips in our maskers was two calls in the 0.625 Hz SAM treatment and fell to one call in the 1.25 and 2.5 Hz SAM treatments; complete calls did not acoustically 'fit' in the shorter dips of maskers modulated at rates faster than 2.5 Hz. Therefore, we might have expected dip listening to occur in the presence of the most slowly fluctuating maskers, as reported for Cope's gray treefrogs and Bow-winged grasshoppers. However, this was not the case (Fig. 3). One possible explanation for the lack of masking release in the 0.625, 1.25, and 2.5 Hz SAM treatments is that one or two consecutive calls without intervening noise are not enough to elicit phonotaxis in green treefrogs. In other words, even in the most slowly fluctuating maskers used in this study, the 'acoustic glimpses' afforded females in these conditions might have been too short to gain any benefit of listening in the dips. As noted earlier, a second, alternative explanation for the lack of masking release in the presence of slowly fluctuating maskers (at least in the 1.25 and 2.5 Hz treatments) is that the maskers competed with the target signal, thereby abolishing or confounding any benefit of listening in the dips. The lack of attraction to the 0.625 Hz masker in experiment 2 (Table 1), however, would be inconsistent with this general interpretation for a lack of dip listening in the presence of slowly fluctuating maskers. Future studies should concentrate on elucidating which of these or other explanations account for the results obtained.

Signal recognition in natural settings

The noise generated in green treefrog choruses is characterized by slow fluctuation rates between about 2 and 7 Hz, with a peak centered at 5 Hz; a second and relatively smaller (<−15 dB) peak was located around 300 Hz (Fig. 2d). The 300-Hz component is likely due to the amplitude modulations in the advertisement call. The predominance of low-frequency fluctuations in green treefrog chorus noise can be attributed to two factors. First,

environmental perturbations on the transmission medium, such as turbulent air, can impose low-frequency fluctuations in ambient noise (Wiley and Richards 1978; Richards and Wiley 1980). Second, the production of repeated and temporally discontinuous calls creates low-frequency modulations that arise from call timing behavior of individuals in a chorus (Nelken et al. 1999). Note, however, that the peak around 5 Hz in the modulation spectrum does not match the rate of approximately 1.8 Hz at which individual males produce advertisement calls. We believe the 5-Hz peak of the modulation spectrum represents vocal interactions among males in a chorus. Green treefrog males actively avoid call overlap with up to two of their nearest neighbors (Höbel and Gerhardt 2007; Jones et al. 2009; Höbel 2011). Given that individual green treefrog males produce advertisement calls about 150 ms long and repeated with a silent interval of approximately 410 ms (560 ms call period), the rate at which bouts of sound are produced by three interacting males that avoid call overlap is approximately 5.3 Hz. Therefore, the peak around 5 Hz in the modulation spectra of green treefrog choruses may be due, in part, to behavioral vocal interactions among males.

Together with our analysis of chorus sounds, our results from experiment 1 suggest that natural level fluctuations in green treefrog chorus noise would have little effect on signal recognition by females for two main reasons. First, we tested fluctuation rates that span the range of predominant level fluctuations in green treefrog choruses and found no evidence of dip listening or modulation masking. Second, studies with human listeners indicate that (i) sinusoidal amplitude modulations provide greater dip-listening benefits than irregular amplitude modulations, and (ii) the amount of masking release and modulation masking increases as a function of increasing modulation depth (Bacon and Grantham 1989; Gustaffson and Arlinger 1994; Bacon et al. 1998). Level fluctuations in chorus settings are not expected to be sinusoidal or to have modulation depths of 100 % (see Fig. 2a–c). Therefore, our SAM maskers potentially offered a 'best case scenario' for observing dip listening and modulation masking compared with natural fluctuations in chorus noise. The fact that neither dip listening nor modulation masking were observed under potentially ideal conditions in the laboratory suggests these processes could have small or even negligible effects on signal recognition in natural settings.

To conclude, we found that female green treefrogs did not experience masking release or modulation masking in the presence of temporally fluctuating backgrounds. The differences between relatively closely related species (i.e., *H. chrysoscelis* and *H. cinerea*) highlight the importance of comparative approaches in behavioral and neurophysiological studies to understand the underlying mechanisms

for communication in noisy environments. Several previous studies have shown the existence of temporally selective neurons and temporal filters in the frog auditory system (Rose and Capranica 1984; Rose et al. 1985; Gooler and Feng 1992; Diekamp and Gerhardt 1995; Alder and Rose 1998; Rose and Gooler 2006). Unfortunately, however, most of these studies have focused solely on how the signal's temporal structure is processed in species with pulsed advertisement calls (but see Klump et al. 2004). Identifying the neurophysiological mechanisms by which a greater diversity of species process amplitude-modulated sounds would help us understand the pattern of results observed in studies of signal recognition in fluctuating noise.

References

Alder TB, Rose GJ (1998) Long-term temporal integration in the anuran auditory system. Nat Neurosci 1:519–523

Arch VS, Narins PM (2008) 'Silent' signals: selective forces acting on ultrasonic communication systems in terrestrial vertebrates. Anim Behav 76:1423–1428

Bacon SP, Grantham DW (1989) Modulation masking: effects of modulation frequency, depth, and phase. J Acoust Soc Am 85:2575–2580

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

Bee MA (2007) Selective phonotaxis by male woodfrogs (*Rana sylvatica*) to the sound of a chorus. Behav Ecol Sociobiol 61:955–966

Bee MA, Micheyl C (2008) The cocktail party problem: what is it? How can it be solved? And why should animal behaviorists study it? J Comp Psychol 122:235–251

Bee MA, Riemersma KK (2008) Does common spatial origin promote the auditory grouping of temporally separated signal elements in grey treefrogs? Anim Behav 76:831–843

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

Bee MA, Buschermöhle M, Klump GM (2007) Detecting modulated signals in modulated noise: (II) neural thresholds in the songbird forebrain. Eur J Neurosci 26:1979–1994

Bee MA, Vélez A, Forester JD (2012) Sound level discrimination by Cope's gray treefrog (*Hyla chrysoscelis*) in the presence and absence of chorus-shaped noise. J Acoust Soc Am 135:4188–4195

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

Brumm H, Slabbekoorn H (2005) Acoustic communication in noise. Adv Stud Behav 35:151–209

Bush SL, Gerhardt HC, Schul J (2002) Pattern recognition and call preferences in treefrogs (Anura: Hylidae): a quantitative analysis using a no-choice paradigm. Anim Behav 63:7–14

Christie K, Schul J, Feng AS (2010) Phonotaxis to male's calls embedded within a chorus by female gray treefrogs, *Hyla versicolor*. J Comp Physiol A 196:569–579

Cohen J (1988) Statistical power analysis for the behavioral sciences, 2nd edn. Erlbaum, Hillsdale

Cooke M (2006) A glimpsing model of speech perception in noise. J Acoust Soc Am 119:1562–1573

Diekamp B, Gerhardt HC (1995) Selective phonotaxis to advertisement calls in the gray treefrog *Hyla versicolor*: behavioral experiments and neurophysiological correlates. J Comp Physiol A 177:173–190

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

Füllgrabe C, Berthommier F, Lorenzi C (2006) Masking release for consonant features in temporally fluctuating background noise. Hear Res 211:74–84

Gerhardt HC (1975) Sound pressure levels and radiation patterns of vocalizations of some North American frogs and toads. J Comp Physiol A 102:1–12

Gerhardt HC (1981) Mating call recognition in the green treefrog (*Hyla cinerea*): importance of two frequency bands as a function of sound pressure level. J Comp Physiol A 144:9–16

Gerhardt HC (1995) Phonotaxis in female frogs and toads: execution and design of experiments. In: Klump GM, Dooling RJ, Fay RR, Stebbins WC (eds) Methods in comparative psychoacoustics. Birkhäuser, Basel, pp 209–220

Gerhardt HC (2001) Acoustic communication in two groups of closely related treefrogs. Adv Stud Behav 30:99–167

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

Gerhardt HC, Klump GM (1988a) Masking of acoustic signals by the chorus background noise in the green treefrog: a limitation on mate choice. Anim Behav 36:1247–1249

Gerhardt HC, Klump GM (1988b) Phonotactic responses and selectivity of barking treefrogs (*Hyla gratiosa*) to chorus sounds. J Comp Physiol A 163:795–802

Gerhardt HC, Tanner SD, Corrigan CM, Walton HC (2000) Female preference functions based on call duration in the gray tree frog (*Hyla versicolor*). Behav Ecol 11:663–669

Gooler DM, Feng AS (1992) Temporal coding in the frog auditory midbrain: the influence of duration and rise-fall time on the processing of complex amplitude-modulated stimuli. J Neurophysiol 67:1–22

Grafe TU, Dobler S, Linsenmair KE (2002) Frogs flee from the sound of fire. P R Soc B 269:999–1003

Gustafsson HA, Arlinger SD (1994) Masking of speech by amplitude-modulated noise. J Acoust Soc Am 95:518–529

Höbel G (2011) Variation in signal timing behavior: implications for male attractiveness and sexual selection. Behav Ecol Sociobiol 65:1283–1294

Höbel G, Gerhardt HC (2007) Sources of selection on signal timing in a treefrog. Ethology 113:973–982

Hofer SB, Klump GM (2003) Within- and across-channel processing in auditory masking: a physiological study in the songbird forebrain. J Neuosci 23:5732–5739

Hulse SH (2002) Auditory scene analysis in animal communication. Adv Stud Behav 31:163–200

Jensen KK (2007) Comodulation detection differences in the hooded crow (Corvus corone cornix), with direct comparison to human subjects. J Acoust Soc Am 121:1783–1789

Jones DL, Jones RL, Ratnam R (2009) Dynamical interactions in a green treefrog chorus. J Acoust Soc Am 126:2270

Joris PX, Schreiner CE, Rees A (2004) Neural processing of amplitude-modulated sounds. Physiol Rev 84:541–577

Klump GM (1996) Bird communication in the noisy world. In: Kroodsma DE, Miller EH (eds) Ecology and evolution of acoustic communication in birds. Cornell University Press, Ithaca, pp 321–338

Klump GM, Langemann U (1995) Comodulation masking release in a songbird. Hear Res 87:157–164

Klump GM, Benedix JH Jr, Gerhardt HC, Narins PM (2004) AM representation in green treefrog auditory nerve fibers: neuroethological implications for pattern recognition and sound localization. J Comp Physiol A 190:1011–1021

Kwon BJ, Turner CW (2001) Consonant identification under maskers with sinusoidal modulation: masking release or modulation interference? J Acoust Soc Am 110:1130–1140

Langemann U, Klump GM (2001) Signal detection in amplitude-modulated maskers. I. Behavioural auditory thresholds in a songbird. Eur J Neurosci 13:1025–1032

Langemann U, Klump GM (2005) Perception and acoustic communication networks. In: McGregor PK (ed) Animal communication networks. Cambridge University Press, Cambridge, pp 451–480

Langemann U, Klump GM (2007) Detecting modulated signals in modulated noise: (i) behavioural auditory thresholds in a songbird. Eur J Neurosci 26:1969–1978

Marshall VT, Schwartz JJ, Gerhardt HC (2006) Effects of heterospecific call overlap on the phonotactic behaviour of grey treefrogs. Anim Behav 72:449–459

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

Nieder A, Klump GM (2001) Signal detection in amplitude-modualted maskers. II. Processing in the songbird's auditory forebrain. Eur J Neurosci 13:1033–1044

Nityananda V, Bee MA (2012) Spatial release from masking in a free-field source identification task by gray treefrogs. Hear Res 285: 86–97

Noldus (2005) EthoVision® Video Tracking System for Automation of Behavioral Experiments: Reference Manual Version 3.1. Noldus Information Technology, Wageningen, The Netherlands

Oldham RS, Gerhardt HC (1975) Behavioral isolation of the treefrogs Hyla cinerea and Hyla gratiosa. Copeia 1975:223–231

Patterson RD, Moore BCJ (1986) Auditory filters and excitation patterns as representations of frequency resolution. In: Moore BCJ (ed) Frequency selectivity in hearing. Academic, London, pp 123–177

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

Ronacher B, 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

Rose GJ, Capranica RR (1984) Processing amplitude-modulated sounds in the auditory midbrain of two species of toads: matched temporal filters. J Comp Physiol A 154:211–219

Rose GJ, Gooler DM (2006) Function of the amphibian central auditory system. In: Narins PM et al (eds) Hearing and sound communication in amphibians. Springer, New York, pp 250–290

Rose GJ, Brenowitz EA, Capranica RR (1985) Species specificity and temperature dependency of temporal processing by the auditory midbrain of two species of treefrogs. J Comp Physiol A 157: 763–769

Rosenthal R, Rosnow RL (1991) Essentials of behavioral research: methods and data analysis, 2nd edn. McGraw-Hill, New York

Schul J, Bush SL (2002) Non-parallel coevolution of sender and receiver in the acoustic communication system of treefrogs. P R Soc B 269:1847–1852

Schwartz JJ, Freeberg TM (2008) Acoustic interaction in animal groups: signaling in noisy and social contexts—introduction. J Comp Psychol 122:231–234

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

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

Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelyd M (1995) Speech recognition with primarily temporal cues. Science 5234: 303–304

Swanson EM, Tekmen SM, Bee MA (2007) Do female anurans exploit inadvertent social information to locate breeding aggregations? Can J Zoolog 85:921–932

Vélez A, Bee MA (2010) Signal recognition by frogs in the presence of temporally fluctuating chorus-shaped noise. Behav Ecol Sociobiol 64:1695–1709

Vélez A, Bee MA (2011) Dip listening and the cocktail party problem in grey treefrogs: signal recognition in temporally fluctuating noise. Anim Behav 82:1319–1327

Vélez A, Bee MA (2012) Signal recognition by green treefrogs (Hyla cinerea) and Cope's gray treefrogs (Hyla chrysoscelis) in naturally fluctuating noise. J Comp Psychol (in press)

Vélez A, Schwartz JJ, Bee MA (in press). Anuran signal perception in noisy environments. In: Brumm H (ed) Animal communication and noise. Springer, New York

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

von Helversen D, von Helversen O (1997) Recognition of sex in the acoustic communication of the grasshopper Chorthippus biguttulus (Orthoptera, Acrididae). J Comp Physiol A 180:373–386

Wiley RH, Richards DG (1978) Physical constraints on acoustic communication in the atmosphere: implications for the evolution of animal vocalizations. Behav Ecol Sociobiol 3:69–94