Studying mate choice using inertial measurement units: A validation study with treefrogs

Saumya Gupta[a] and Mark A. Bee[a, b,*]

[a] Department of Ecology, Evolution, and Behavior, University of Minnesota - Twin Cities, St. Paul, MN, U.S.A.

[b] Graduate Program in Neuroscience, University of Minnesota - Twin Cities, Minneapolis, MN, U.S.A.

* Correspondence: Mark Bee, Department of Ecology, Evolution, and Behavior, University of Minnesota - Twin Cities, 1479 Gortner Ave Ste 140, St. Paul, MN, 55108 U.S.A.

Email addresses: saumya1692@gmail.com (S. Gupta), mbee@umn.edu (M. Bee

Investigations of mate choice continue to address fundamental questions about the mechanisms and evolution of animal behaviour. A common behavioural assay used to study acoustically guided mate choice with playback experiments is phonotaxis, a typically robust response in which a chooser approaches acoustic signals, such as courtship songs or mating calls. Robust empirical studies of phonotaxis often require substantial laboratory facilities, such as a dedicated and sound-treated room or enclosure, in which the acoustic environment is controlled and in which animals are freely able to move about. The financial and space resources required to outfit a research laboratory to investigate phonotaxis may be sufficiently prohibitive such that some researchers are excluded from undertaking bioacoustic behavioural research. Here, we validate a new device designed to measure animal movements related to phonotaxis using an inertial measurement unit (IMU). The device is small and portable; it can be constructed for less than $300 US dollars; and the build instructions and code for operation are freely available. In a series of four experiments with treefrogs, we demonstrate using the device that an IMU-based approach to measuring animal movement can replicate a broad range of findings from traditional phonotaxis experiments on species recognition and sexual selection. We conclude by discussing several possible uses for IMU-based measurements of phonotaxis.

Keywords
Acoustic communication
Advertisement calls
Anuran
Auditory perception
CRoAK
Gray treefrog
Mating behaviour
Sensory biology
Tree frog
Vocal communication

Elucidating the mechanisms, evolution, and consequences of mate choice remains a fundamental goal of research on animal behaviour (Rosenthal, 2017). The empirical methods and behavioural assays used to investigate the mate preferences that “choosers” (most commonly females) exhibit when selecting particular “courters” (most commonly males) are diverse and vary by taxonomic group (Rosenthal, 2017). In fish, for example, female preferences are typically assessed using measures of time spent near stimulus males (Morris, Nicoletto, & Hesselman, 2003; Wong, So, & Cummings, 2011) or video depictions of males (Morris et al., 2003; Morris, Rios-Cardenas, & Tudor, 2006; Rosenthal & Evans, 1998). In songbirds, studies of acoustically mediated mate preferences commonly rely on measuring the copulation solicitation displays that females produce in response to hearing attractive male songs (Draganoiu, Nagle, & Kreutzer, 2002; O'Loghlen & Beecher, 1999). In contrast, mate preferences in acoustically signalling orthopteran insects and anuran amphibians are commonly assessed using female phonotaxis toward speakers broadcasting real or synthetic male signals (Gerhardt, 1995; Gerhardt & Huber, 2002).

Conducting phonotaxis tests in the laboratory can present challenges that exclude some researchers from studying acoustically mediated mate preferences. Phonotaxis tests that simulate signalling males spaced at realistic distances can require testing areas that are quite large relative to the size of the animal. Past studies, for example, have used a 3 m long rectangular arena for testing crickets (Bailey & Zuk, 2008), a 3 m × 3 m square arena for testing katydids (Snedden & Greenfield, 1998), or a 2 m diameter circular arena (Bush, Gerhardt, & Schul, 2002) or an artificial, octagonal pond (2 m on a side) enclosed in a greenhouse (Schwartz, Buchanan, & Gerhardt, 2001) for testing frogs. Conducting such tests under quiet and well controlled acoustic conditions requires enclosed spaces, such as semi-anechoic sound chambers (e.g., Bee & Schwartz, 2009; Reichert, Symes, & Höbel, 2016) or sound-treated rooms (e.g., Heinen-Kay, Strub, & Zuk, 2018; Snedden & Greenfield, 1998), that provide adequate transmission loss (to reduce extraneous noise) and in which hard surfaces, such as the walls and ceiling, can be treated with acoustic foam (to reduce reverberations). The price of a high-quality semi-anechoic sound chamber large enough to conduct phonotaxis experiments can easily exceed $80,000 US dollars. In many academic settings, the intersection of needing to control the acoustic environment in a large, enclosed space devoted to animal behaviour studies can create substantial barriers to research due to limited financial resources and constraints on available laboratory space. Alternative methods of assessing chooser phonotaxis behaviour could help remove these barriers.

In this study of Cope's grey treefrog (*Hyla chrysoscelis*) and the eastern grey treefrog (*Hyla versicolor*), we investigated an alternative method for measuring phonotaxis behaviour that is less demanding of financial and space resources than traditional laboratory phonotaxis tests. Grey treefrogs represent a cryptic species complex that have been well studied in the context of mate preferences using traditional phonotaxis methods (Gerhardt, 2001; Gerhardt & Huber, 2002). Gravid female treefrogs commonly exhibit phonotaxis in response to hearing the advertisement calls of a potential mate using short hops or bouts of walking as they approach the sound source. In grey treefrogs, phonotaxis movements typically occur shortly after the offset of a call (Beckers & Schul, 2004; Gupta, Marchetto, & Bee, 2020). We recorded these phonotaxis-related movements using the Customizable Recorder of Animal Kinesis (CRoAK; Gupta et al., 2020), a small, portable apparatus that can be built for under $300 USD to measure phonotaxis in frogs and other small animals (Fig. 1; see Supplementary Material). CRoAK consists of an Arduino-controlled, 9 degree-of-freedom (9DoF) inertial measurement unit (IMU) that senses movements of the animal in a mesh enclosure as it listens and responds to acoustic stimuli. The unit is compact enough to be used in almost any laboratory setting (and potentially also in the field), and it can easily be installed in bench-top acoustic chambers, which typically cost a small fraction of larger sound chambers. Gupta et al. (2020) provided full details on how to construct and operate a CRoAK device and demonstrated how CRoAK's IMU continuously registers the movements made by gravid female treefrogs in response to hearing repeated advertisement calls simulating a calling male. Here, in a series of four experiments, we validate CRoAK for use in studies of mate preferences by demonstrating that the device can reproduce results from traditional phonotaxis tests that investigated a range of questions about species recognition and sexual selection in frogs.

METHODS

Ethical Note

The procedures described here met all legal requirements of the United States of America, where this work was performed, were approved by the University of Minnesota Institutional Animal Care and Use Committee (IACUC protocol 1701-34456A and 2001-37746A) and adhered to the ASAB/ABS Guidelines for the ethical treatment of animals. Research permits and special use permits to collect frogs from publicly owned lands were issued by the Minnesota Department of Natural Resources, the Ramsey County Department of Parks and Recreation and the Three Rivers Park District.

Subjects

General protocols for collecting and handling female grey treefrogs and for reducing stress associated with collecting and handling closely followed those described in previous studies (Bee, 2007, 2008a, 2008b; Bee & Riemersma, 2008; Bee & Swanson, 2007; Gerhardt, 1995; LaBarbera, Nelson, & Bee, 2020; Ward et al., 2013b). We collected 109 pairs of grey treefrogs in amplexus during the breeding seasons (May-July) of 2019 and 2020. Collections were made at night, between 2200 and 0100 h from ponds in the Carver Park Reserve (Carver County, MN, U.S.A.) and the Tamarack Nature Center (Ramsey County, MN, U.S.A.). Pairs were collected gently by hand; placed as a pair in a small (710 ml), clear plastic container fitted with a plastic lid equipped with holes to facilitate gas exchange; and transported (usually within 3 hrs) to the laboratory, where the containers were filled to a depth of approximately 1 cm with clean, aged tap water to facilitate cutaneous respiration. Most subjects (N = 88; 81%) were released with their mate at their location of capture within 24 to 48 h. We did not individually mark these subjects because female grey treefrogs probably only produce one or at most two egg clutches per year in Minnesota (Ritke, Babb, & Ritke, 1990); most animal collections for a single experiment were completed in a single year; and collections were made from multiple ponds in two different natural areas, each with large populations of grey treefrogs. Therefore, we estimated a low likelihood of unknowingly recapturing the same subject and retesting it in the same experiment. In 2019, a group of females (N = 21, 19%) that had previously oviposited and whose mates were returned to their location of capture were kept in the laboratory to measure the behavioural responses of nongravid females outside the breeding season. These females were tested in August 2019, approximately one month after the end of the local breeding season during which they were collected. We were not permitted to return these animals to the field; therefore, they were euthanized by overdose of MS-222 after reaching the experimental endpoint, as required by our IACUC and permitting agency.

Prior to testing during the breeding season, pairs were housed in their collection containers at 4° C to delay egg laying by the female and to prolong the period of behavioural responsiveness (Gerhardt, 1995). Temporary housing at cold temperatures does not elevate endogenous corticosterone levels, which would generally be expected to increase if the animals exhibited a stress response (Gall, Bee, & Baugh, 2019). Subjects tested after the breeding season were housed under a 12L:12D light cycle in groups of up to five animals in plastic terraria (35 cm × 22 cm × 25 cm) equipped with perches or refugia and a glass dish filled daily with clean, aged water. These subjects were fed vitamin dusted crickets two to three times per week.

Apparatus and Testing Protocol

Here, we briefly describe how the CRoAK unit works and the testing protocols used to validate its utility for testing hypotheses about acoustically guided mating behaviours in frogs. Additional details on design and construction are reported elsewhere (Gupta et al., 2020). Figure 1 schematically depicts the CRoAK unit (see the Supplementary Material for an animated 3D rendering). The unit consisted of a small enclosure (86 mm diameter, 60 mm height) made from plastic canvas and sprouting lids. The walls were glued into the lower sprouting lid; the upper sprouting lid could be removed for placing animals inside. The enclosure was sufficiently large to allow female grey treefrogs to freely move around inside. The enclosure was suspended by thread from a steel bar that was attached to a suspension spring mounted from a shelf bracket, which was bolted to a plywood board for support (Fig. 1). The bottom of the enclosure was attached by thread to the plywood frame to limit the movement of the enclosure when animals responded. This was necessary to suppress sustained and excessive oscillations and swinging of the enclosure when animals moved, but still allowed enough movement of the enclosure to be registered by the IMU.

A 9DoF IMU (SEN-13944, SparkFun Electronics, Boulder, CO, U.S.A.) was attached to the bottom of the enclosure. This IMU integrates an accelerometer (linear acceleration), a gyroscope (angular velocity), and a magnetometer (direction of the magnetic field), each of which could record data in three axes (x, y and z). In the implementation of CRoAK used for the experiments reported here, we relied solely on angular velocity data recorded by the gyroscope to quantify the motion of the enclosure (Gupta et al., 2020). A sound detector (SEN-12642, SparkFun Electronics, Boulder, CO, U.S.A.) was located near the unit to record the amplitude envelope of acoustic stimuli. The IMU and sound detector were controlled by an Arduino Nano microcontroller (Arduino LLC, Boston, MA, U.S.A.). The source code for reading sensor data from the IMU and sound detector is reported in Gupta et al. (2020). We simultaneously captured continuous output from the IMU and sound detector using SerialPlot v0.12 (~50 samples/s; written by Hasan Yavuz Özderya) running on a Dell OptiPlex 5060 (Dell Computer Corporation, Round Rock, TX, U.S.A.).

The CRoAK unit was mounted on a plywood frame located on the floor of a portable anechoic chamber (51 cm × 61 cm × 60 cm inside dimensions; Eckel Industries, Cambridge, MA, U.S.A.). The inside walls, door and a portion of the floor of the chamber were lined with Sonex acoustic foam panels (Model VLW-60; Pinta Acoustic, Inc. Minneapolis, MN, U.S.A.) to reduce reverberation. A speaker (TEBM65C20F-8 Balanced-Mode Radiator, Tectonic Audio

Labs, Inc., Woodinville, WA, U.S.A.) was mounted to and elevated above the floor of the chamber just above the height of the enclosure at a distance of 15 cm from the enclosure's centre. Acoustic stimuli consisted of synthetic advertisement calls generated in MATLAB v2018b. Stimulus calls were output from the soundcard of the Dell PC using Adobe Audition 3.0 (Adobe Inc, San Jose, CA, U.S.A.) and amplified using a Crown XLS 1502 amplifier (Harman Professional, Northridge, CA, U.S.A.). Sound pressure levels (dB SPL re 20 μPa) were calibrated by placing the microphone of a Larsen Davis 831 sound level meter (Depew, NY, U.S.A.) at the approximate position of a subject's head inside the enclosure.

At the time of testing, subjects were allowed to acclimate at 20 ± 1° C inside a temperature-controlled incubator for at least 30 min before commencing the first test, and they were maintained at this temperature throughout testing. To begin an experimental trial, a subject was removed from the incubator and placed in the CRoAK enclosure. Each subject typically heard up to 10 stimulus sequences during an experimental trial, which always lasted 30 min or less. Between presentations of different stimulus sequences within a trial, subjects were placed in the incubator for a time-out of at least 2 min before being returned to the enclosure before presentation of the next stimulus sequence. There is little evidence for "carryover" effects between consecutive phonotaxis tests of frogs, even in frogs retested after short time intervals (Akre & Ryan, 2010; Gerhardt, 1981; Gerhardt, Tanner, Corrigan, & Walton, 2000). Prior to the start of each stimulus sequence, subjects were allowed to acclimate in the enclosure for 30 s. Each stimulus sequence consisted of broadcasting a series of between 10 and 30 synthetic stimulus calls, the acoustic properties of which differed and were dictated by each experiment. Within a stimulus sequence, stimulus calls repeated with a period of 10 s, which approximates a natural calling rate in grey treefrogs. In all experiments, one stimulus was designated as a 'standard call' and had properties approximating average values in local populations. We refer to the complete set of stimulus calls occurring across all stimulus sequences in an experiment as the 'stimulus set' for that experiment. For statistical analyses, responses to the standard call were used for normalizing responses to all other stimuli in a stimulus set. Subjects were not tested in more than one experimental trial.

Data Processing and Analysis

Raw measurements of angular velocity (ω) were recorded for three rotational axes (x, y and z) in units of deg/s and subsequently converted to the magnitude of angular velocity according to equation 1:

Eq. 1 $$\left(\underline{\omega} = \sqrt{\omega_x^2 + \omega_y^2 + \omega_z^2}\right).$$

Using MATLAB v2018b, we computed the area under the curve in plots of $\underline{\omega}$ versus time for the 1 s that immediately preceded the onset of each stimulus call in a stimulus sequence and the 1 s that immediately followed the offset of the stimulus call. For each stimulus call, we then computed the difference in these two areas as a measure of the animal's response to that call. We computed each subject's overall response to each different type of stimulus call in the stimulus set as the median of the area difference across all broadcasts of the stimulus call across all stimulus sequences in the experiment. This procedure allowed us to compute the magnitude of the typical movement response of each subject that was evoked by a particular stimulus call relative to their own level of movement occurring just prior to stimulation by that same call. Subjects typically exhibited little movement except when responding to a stimulus call. To normalize responses across subjects in a given experiment and across the different stimulus calls in a stimulus set, we divided each subject's median response to each stimulus call, including the standard call, by the sample average of median response to the standard call in that experiment and expressed these responses as a percentage. This normalization procedure allowed us to compare movement responses to each stimulus call in a stimulus set relative to the magnitude of movement evoked by the standard call that was included in the same stimulus set. Note that normalized responses can be negative if the IMU registered more activity immediately prior to a stimulus call relative to immediately after it.

We used Generalized Linear Mixed-effects Models (GLMM; package *lme4*) in R 4.0.2 (Bates, Mächler, Bolker, & Walker, 2015; R Core Team, 2020) to compare normalized responses across the different stimulus calls in the stimulus set unique to each experiment. Frog ID was included as a random effect in all models. The predictor variables included in each GLMM model differed depending on the specific experiment and are indicated below. The α level for each experiment was corrected for the number of multiple comparisons in that experiment using the Holm's sequential Bonferroni procedure.

EXPERIMENT 1 – SPECIES DISCRIMINATION

Methods

Gravid female grey treefrogs exhibit selective phonotaxis toward calls produced by a conspecific male, whereas nongravid females tested outside of their breeding season or immediately following oviposition are not responsive to calls (Gall et al., 2019; Gerhardt, 2001; Gordon & Gerhardt, 2009; Ward et al., 2015). In this experiment with *H. chrysoscelis*, we validated CRoAK

by measuring the species specificity of phonotaxis behaviour and its dependence on reproductive state. We presented subjects with conspecific and heterospecific calls at two different times of the year, during the breeding season when subjects were reproductively ready and during the non-breeding season after the subjects had oviposited.

The stimulus set consisted of a synthetic *H. chrysoscelis* call and a synthetic *H. versicolor* call, each modelled after an advertisement call having acoustic properties close to the average for that species recorded in local populations and adjusted to 20° C (Gupta, Alluri, Rose, & Bee, 2021; Ward et al., 2013b). The *H. chrysoscelis* call consisted of a train of 30 pulses (50 pulses/s, 20-ms pulse period, 10-ms pulse duration, 3.1-ms inverse exponential rise time and 5.4-ms exponential fall time). Each pulse was composed of two harmonically related, phase-locked sinusoids having frequencies (and relative amplitudes) of 1.25 kHz (-11 dB) and 2.5 kHz (0 dB). The *H. versicolor* call consisted of 18 pulses (16.7 pulses/s, 60-ms pulse period, 30-ms pulse duration, 20-ms linear rise time and 10-ms linear fall time), with each pulse composed of two harmonically related, phase-locked sinusoids having frequencies (and relative amplitudes) of 1.2 kHz (-5 dB) and 2.4 kHz (0 dB). Each stimulus sequence consisted of repeating both synthetic calls 15 times each in random order (30 calls per sequence). All subjects were tested with stimulus sequences in which each stimulus call was broadcast at 85 dB SPL at the position of the subject. This signal level approximates the amplitude of a male calling at a distance of 1 m (Gerhardt, 1975). Stimulus sequences were also presented at two additional signal levels (65 dB and 75 dB SPL) to those subjects tested during the breeding season, with the order of signal level determined randomly. The *H. chrysoscelis* call presented at 85 dB SPL during the breeding season was the standard call for normalizing responses.

Results and Discussion

Females tested during the breeding season responded strongly to conspecific calls but not heterospecific calls. Figure 2 shows a representative trace of an IMU recording of responses to two consecutive stimuli in a sequence, a heterospecific call (*H. versicolor*) followed 10 s later by a conspecific call (*H. chrysoscelis*). The IMU registered no obvious difference in movements in the 1 s time analysis windows immediately prior to and following the heterospecific call. In contrast, the IMU recorded a response following the conspecific call that was much larger than the movement recorded 1 s prior to the conspecific call. The difference between the areas under the recorded IMU curve in the 1 s immediately prior to and following a stimulus call determined the magnitude of response to that call.

Normalized responses to conspecific and heterospecific calls are shown in Figure 3. To statistically examine the species-specificity of evoked movement responses, we conducted a GLMM comparing the normalized responses of females as functions of the species identity of the stimulus (conspecific versus heterospecific), the signal level (65, 75, 85 dB SPL), and their interaction. An ANOVA comparing models with and without the interaction term revealed a significant interaction between species identity and signal level (χ^2 = 16.9, df = 2, P = 0.0002). By definition, the mean normalized response to the conspecific call at 85 dB (i.e., the standard call) was 100% (Fig. 3a). The mean normalized responses to the conspecific call at 75 dB and 65 dB were 48.1% and 30.5%, respectively (Fig. 3a). At all signal levels, the 95% confidence intervals for responses to conspecific calls excluded 0%. By comparison, the mean normalized responses to the heterospecific (H. versicolor) call broadcast at 85 dB, 75 dB, and 65 dB SPL were -4.6%, -0.8%, and -5.6%, respectively. At all signal levels, the 95% confidence intervals for responses to the heterospecific call included 0% (Fig. 3a), indicating no reliable detection of movements in response to heterospecific calls at all signal levels. Responses to conspecific calls were significantly higher than responses to heterospecific calls at all signal levels tested (65 dB: t = 2.93, df = 110, *P* = 0.004; 75 dB: t = 3.97, df = 110, P < 0.001; 85 dB: t = 8.45, df = 110, *P* < 0.001).

Post-oviposition females tested outside the breeding season did not respond to calls (Fig. 3b). The mean response of non-breeding females was -0.8% to both conspecific and heterospecific calls presented at 85 dB (Fig. 3b). The 95% confidence intervals for responses to conspecific calls by females in non-breeding condition included 0% (Fig. 3b). Responses to the conspecific call presented at 85 dB were significantly lower in non-breeding females compared with those tested during the breeding season (t = -5.83, df = 40, *P* < 0.0001). There was no difference in the magnitude of response evoked by conspecific and heterospecific calls outside the breeding season (t = -0.02, df = 40, *P* = 0.9850) (Fig. 3b).

The results from Experiment 1 confirm that an IMU can reliably record the acoustically evoked phonotaxis movements of gravid female treefrogs, and that IMU recordings can be used effectively to measure species discrimination behaviour. These results replicate findings from numerous previous studies of species recognition and reproductive isolation in grey treefrogs conducted using traditional phonotaxis methods (Bee, 2008a; Gerhardt & Doherty, 1988; Gerhardt, Dyson, Tanner, & Murphy, 1994; Li, Schrode, & Bee, 2022; Littlejohn, Fouquette, & Johnson, 1960). Our results further demonstrate that response magnitudes are higher at higher signal amplitudes, a result also known from previous phonotaxis studies of grey treefrogs (Beckers & Schul, 2004; Bee & Schwartz, 2009; see also Experiment 4). In addition, previous

phonotaxis studies of grey treefrogs have shown that approach toward sources of conspecific calls depends on the subject's reproductive state: nongravid females tested outside the breeding season or immediately after oviposition during the breeding season fail to respond to conspecific calls (Gall et al., 2019; Gordon & Gerhardt, 2009; Ward et al., 2015; see also Littlejohn et al., 1960). We replicated this effect of variation in reproductive state: gravid females tested in the breeding season were highly selective for conspecific calls but those tested outside of the normal breeding season approximately one month after oviposition were entirely unresponsive to both conspecific and heterospecific calls (Fig. 3b). Together, results from Experiment 1 confirm that IMUs can be used successfully to investigate mate preferences in the context of species discrimination.

EXPERIMENT 2 – PULSE-RATE PREFERENCE FUNCTIONS

Methods

Pulse rate is a species recognition cue in grey treefrogs (Bush et al., 2002; Gerhardt & Doherty, 1988; Schul & Bush, 2002). In traditional phonotaxis tests, females maximally respond to pulse rates near or slightly higher than the population mean for conspecific calls and show decreased responses to calls that have lower or higher pulse rates (e.g., Bush et al., 2002). In this experiment with subjects of both *H. chrysoscelis* and *H. versicolor*, we validated CRoAK by measuring the pulse rate preference functions of the two species. The stimulus set for each species consisted of seven stimulus calls differing in pulse rate (measured in pulses/s), one of which was a standard call having a pulse rate close to the population average for that species. Following Bush et al. (2002), we examined the responses of *H. chrysoscelis* subjects to calls with pulse rates of 20, 30, 40, 50 (standard call), 60, 80, and 125 pulses/s, and of *H. versicolor* subjects to calls with pulse rates of 5, 8, 10, 20 (standard call), 30, 50, and 100 pulses/s. For both species, the range of pulse rates used encompassed the range of natural variation in pulse rate at 20° C recorded in local populations (Gupta et al., 2021; Ward et al., 2013b). The number of pulses per call was adjusted by adding or removing the appropriate number of pulses (and interpulse intervals) to maintain a nearly constant call duration (in ms) across stimuli (as in Bush et al., 2002). The pulse duration and interpulse intervals of calls were manipulated to obtain the desired pulse rates, but otherwise, all calls had the same acoustic properties as the standard call of the appropriate species.

Each stimulus sequence consisted of 10 stimulus calls. Four of the stimulus calls in a sequence – the first, fourth, seventh, and last – were the standard call. The remaining six

stimulus calls each had a different pulse rate chosen from within the stimulus set and were interspersed in random order within a stimulus sequence, with two stimulus calls differing in pulse rate presented between each repetition of the standard call. A total of 10 stimulus sequences with different randomized stimulus orders were presented to each subject. All calls were played back at a signal level of 85 dB SPL measured at the position of the subject.

Results and Discussion

Females of both *H. chrysoscelis* and *H. versicolor* exhibited species-specific patterns of pulse-rate selectivity (Fig. 4). We found significant overall effects of pulse rate on normalized responses for both species (*H. chrysoscelis*: χ^2 = 78.6, df = 6, P < 0.0001; *H. versicolor*: χ^2 = 52.2, df = 6, P < 0.0001). As expected, responses were generally maximal near the species-specific pulse rates in local populations (~50 pulses/s for *H. chrysoscelis*; ~21 pulses/s for *H. versicolor*) and decreased at higher and lower pulse rates (Fig. 4). Compared with responses to the 50 pulses/s standard call in *H. chrysoscelis* (Fig. 4, filled symbols), responses were significantly lower at pulse rates of 20 pulses/s (t = -7.1, df = 126, P < 0.0001), 30 pulses/s (t = -5.4, df = 126, P < 0.0001), 40 pulses/s (t = -4.6, df = 126, P < 0.0001) and 125 pulses/s (t = -5.9, df = 126, P < 0.0001), but not 60 pulses/s (t = 0.4, df = 126, P = 0.686) and 80 pulses/s (t = -1.6, df = 126, P = 0.103). For *H. versicolor* (Fig. 4, open symbols), responses were significantly lower than those evoked by the 20 pulses/s standard call at pulse rates of 5 pulses/s (t = -5.3, df = 138, P < 0.0001), 8 pulses/s (t = -4.9, df = 138, P < 0.0001), 10 pulses/s (t = -4.2, df = 138, P < 0.0001), 50 pulses/s (t = -4.0, df = 138, P < 0.0001) and 100 pulses/s (t = -5.5, df = 138, P < 0.0001), but not 30 pulses/s (t = -0.40, df = 138, P = 0.69).

In tests of both *H. chrysoscelis* and *H. versicolor,* the results from Experiment 2 confirm that an IMU can reliably measure preference functions for species recognition cues. Moreover, use of an IMU was able to replicate both the general shape of, and species differences in, the pulse-rate preferences functions of females in this cryptic species group. A key finding from previous studies of mate choice in grey treefrogs is that females attend to information related to the timing of pulses in advertisement calls to identify conspecific males (Bush et al., 2002; Gerhardt, 1991, 2001, 2008; Gerhardt & Doherty, 1988; Gupta et al., 2021; Hanson et al., 2016; Rose, Brenowitz, & Capranica, 1985; Rose et al., 2015; Schul & Bush, 2002). In both species, pulse rate appears to be under stabilizing or weakly directional selection and is considered a static call property because it varies little within males (Gerhardt, 1991; Gerhardt & Brooks, 2009; Tanner, Ward, Shaw, & Bee, 2017; Ward, Buerkle, & Bee, 2013a). Bush et al. (2002) measured pulse-rate preference functions for both grey treefrog species using traditional

phonotaxis methods based on single-stimulus tests. Females of both species had peak preferences near the population mean pulse rate of calls produced by males of their own species, and they responded less strongly or not at all to faster and slower pulse rates. Neurophysiological recordings from the midbrain of both grey treefrog species have revealed populations of neurons in the inferior colliculus with pulse-rate selectivity that broadly mirrors behavioural preference functions for pulse rate (Diekamp & Gerhardt, 1995; Gupta et al., 2021; Rose et al., 1985, 2015). The pulse-rate preference functions we measured in Experiment 2 are strikingly similar to those reported by Bush et al. (2002; cf. our Fig. 4 with their Fig. 3a). These results extend those of Experiment 1 by showing that IMUs can be used successfully to investigate the specific properties of acoustic signals to which females attend when exercising mate choice preferences for conspecific males.

EXPERIMENT 3 – CALL DURATION PREFERENCE FUNCTION

Methods

In animals that produce pulsatile sounds, it is common for females to prefer longer signals with more pulses in the context of intraspecific mate choice (Ryan & Keddy-Hector, 1992). Male grey treefrogs differ on average in how many pulses they incorporate into each call, and females exhibit robust, directional preferences for males that produce longer calls with more pulses (Bee, 2008b; Gerhardt et al., 2000; LaBarbera et al., 2020; Ward et al., 2013b; Welch, Semlitsch, & Gerhardt, 1998). Moreover, the directional preference for call duration is nonlinear in that females discriminate more strongly against shorter-than-average calls (Bee, 2008b; Gerhardt et al., 2000; LaBarbera et al., 2020; Schwartz et al., 2001; Schwartz, Huth, & Hutchin, 2004). In this experiment with *H. chrysoscelis*, we validated CRoAK by measuring a preference function for call duration. The stimulus set consisted of stimulus calls based on the *H. chrysoscelis* standard call but having either 5, 10, 20, 30 (standard call), or 40 pulses per stimulus call. Five stimulus sequences were created, with each consisting of 15 presentations of the same stimulus call having one of the five specified number of pulses per call. Subjects were presented with the five stimulus sequences in random order. Aside from pulse number, and hence call duration, the acoustic properties of all calls were the same as the *H. chrysoscelis* standard call. All stimuli were presented at a signal level of 85 dB SPL measured at the position of the subject.

Results and Discussion

Females exhibited stronger and nonlinear responses to longer calls having more pulses. As shown in Figure 5, there was a sigmoidal increase in response magnitude as a function of increasing pulse number. Subjects failed to respond to the 5-pulse call, but the 10-pulse call was able to elicit a low level of response. Responses increased monotonically between 10 pulses/call and 30 pulses/call, and then increased more slowly above 30 pulses/call. There was a significant effect of pulse number on normalized responses (χ^2 = 66.7, df = 4, P < 0.0001). Compared with the standard 30-pulse call, normalized responses were lower for the 5-pulse (t = -6.9, df = 84, P < 0.0001), 10-pulse (t = -6.2, df = 84, P < 0.0001), and 20-pulse (t = -3.2, df = 84, P = 0.0018) calls. There was no significant difference in response to the 30-pulse and 40-pulse calls (t = 1.1, df = 84, P = 0.2676), though responses to the latter were slightly greater (Fig. 5). The population mean and standard deviation (SD) for call duration is approximately 30 ± 4 pulses/call (Ward et al., 2013b). Thus, 20-pulse and 40-pulse calls deviated from the standard call (30 pulses/call) by -2.5 SD and +2.5 SD, respectively. The difference in normalized response between the 20-pulse and 30-pulse calls was close to 50%, whereas that between the 30-pulse and the 40-pulse call was less than 25%, indicating a nonlinear scaling of response strength to variation in pulse number.

Results from Experiment 3 confirm that an IMU is able not only to measure directional preference functions for sexually selected traits, but also to measure distinctive features of those preference functions, such as nonlinearity. These results replicate findings from previous studies that have shown using traditional phonotaxis methods that females of both grey treefrog species prefer males that produce longer calls with more pulses (Bee, 2008b; Gerhardt, 1991; Gerhardt & Brooks, 2009; Gerhardt, Dyson, & Tanner, 1996; Gerhardt et al., 2000; LaBarbera et al., 2020; Schwartz et al., 2001; Schwartz et al., 2004; Tanner et al., 2017; Ward et al., 2013b). In *H. versicolor*, males that produce longer calls sire offspring of higher fitness (Welch et al., 1998), and call duration can be moderately heritable (h^2 = 0.35 to 0.41)(Welch, Smith, & Gerhardt, 2014). Females of both species discriminate most strongly against males that produce shorter-than average calls (Bee, 2008b; Bush et al., 2002; Gerhardt et al., 2000; LaBarbera et al., 2020; Schwartz et al., 2001). Using a series of single-stimulus tests based on a traditional phonotaxis approach, Bush et al. (2002; see their Fig. 5) found that females were generally unresponsive to the shortest calls tested and that the strength of phonotaxis responses increased toward an asymptotic level as pulse number increased. Such a preference function showing diminishing returns of increasing pulse number is strikingly similar to the preference function we found using an IMU to measure behavioural responses (cf. our Fig. 5 with Fig. 5 in Bush et al. 2002). In particular, our finding that responses to the 20-pulse call, but

not the 40-pulse call, differed significantly from the average (30-pulse) call are consistent with the known nonlinearity in female preferences for call duration in *H. chrysoscelis* (Bee, 2008b; LaBarbera et al., 2020). Thus, results from Experiment 3 extend those of Experiments 1 and 2 by showing that the IMU implemented in CRoAK can be used effectively to investigate intraspecific mate preferences in frogs. Our IMU results additionally suggest that females responded, albeit at a very low level, when calls had 10 pulses but not when there were only 5 pulses/call. This finding is consistent with results from traditional phonotaxis tests showing that, on average, calls with about 8 pulses are required to elicit positive phonotaxis, and that this behavioural pulse-number threshold was very similar to neural pulse-number thresholds in the auditory midbrain (Gupta et al., 2021).

EXPERIMENT 4 – CALL AMPLITUDE PREFERENCE FUNCTION

Methods

In the context of intraspecific mate choice in acoustically signalling animals, females also commonly prefer males that produce signals of greater amplitude (Ryan & Keddy-Hector, 1992). In grey treefrogs, males differ in call amplitude (Gerhardt, 1975), and females prefer males that produce higher-amplitude calls (Bee, Vélez, & Forester, 2012; Fellers, 1979). Preference functions for call amplitude in grey treefrogs often resemble traditional psychometric functions, which relate subject response to stimulus amplitude. Averaged across females, they have a sigmoid shape reflective of low responsiveness at low amplitudes, increasing responsiveness at increasing intermediate amplitudes, and a saturation of, or even decrease in, responsiveness at the highest amplitudes (Beckers & Schul, 2004; Bee & Schwartz, 2009). In this experiment with *H. chrysoscelis*, we validated CRoAK by measuring call amplitude preference functions. The stimulus set consisted of presenting the *H. chrysoscelis* call from Experiment 1 at each of 11 sound pressure levels in 6-dB steps ranging between 25 dB and 85 dB SPL at the position of the subject. The stimulus set additionally included a ‘no signal’ sham treatment in which silence was inserted in place of a stimulus call. A total of 10 stimulus sequences was presented to each subject, and within each stimulus sequence the order of the 11 stimulus calls and the sham treatment was randomly determined. The stimulus call presented at 85 dB SPL was designated as the standard call for this experiment.

Results and Discussion

Females of *H. chrysoscelis* exhibited increasingly strong responses as call amplitude increased (Fig. 6). There was a significant overall effect of call amplitude on normalized responses (χ^2 =

251.6, df = 11, $P < 0.0001$). The call amplitude preference function (i.e., the population-level psychometric function) increased monotonically between approximately 55 dB and 85 dB (Fig. 6). The mean normalized responses were above the upper bound of the 95% confidence interval around the mean normalized response in the sham trial (dashed line in Fig. 6) at call amplitudes of 55 dB and above. Compared with the no-signal sham trials, normalized responses were significantly higher at signal levels of 67 dB and above after corrections for multiple comparisons (ts > 4.4, df = 253, Ps < 0.0001). Normalized responses at 61 dB were also higher than those from the sham trials, but this difference was not quite significant after corrections for multiple comparison (t = 2.4, df = 253, $P = 0.0155$). Compared with responses at 85 dB, normalized responses were significantly lower at signal levels of 73 dB and lower (ts < -6.41, df = 253, Ps < 0.0001).

Results from Experiment 4 replicate the general pattern expected for a psychometric function relating response level to call amplitude as well as patterns demonstrated in previous studies of this species that used traditional phonotaxis tests (Beckers & Schul, 2004; Bee & Schwartz, 2009; Bee & Swanson, 2007). In previous phonotaxis studies, estimates of behavioural thresholds for responding to conspecific calls with phonotaxis have ranged between about 35 dB and 43 dB SPL (Bee & Schwartz, 2009; Nityananda & Bee, 2012), with substantial variation among individuals (e.g., 29 dB to 55 dB SPL; Nityananda & Bee, 2012). In the present study, statistical differences from baseline levels in the no-signal sham condition first appeared at a signal level of 67 dB SPL after corrections for multiple comparisons, although there was suggestive visual evidence that behavioural responses had diverged from baseline at somewhat lower signal levels in the range of 49 dB to 61 dB SPL. Alternative analysis methods not reported here, such as those based on more standard threshold measures (e.g., Sakitt, 1973), yielded quantitatively similar threshold estimates. Thus, measures of behavioural responsiveness as a function of call amplitude may be somewhat less sensitive than more traditional phonotaxis methods at measuring the lowest call amplitude that will elicit phonotaxis. It should be borne in mind, however, that traditional phonotaxis tests allow unrestricted movement within the sound field, thereby creating potential opportunities for random movements from a starting position toward a sound source to increase the received call amplitude above an individual's behavioural threshold, and thus leading to artificially low threshold estimates.

GENERAL DISCUSSION

The main result of this study is that using an IMU to record phonotaxis movements yielded results that were qualitatively – and in some cases quantitatively – similar to those obtained using more traditional phonotaxis tests conducted in large anechoic sound chambers. These results confirm that IMU-based devices could significantly broaden the use of animal phonotaxis as a means of testing key evolutionary, physiological, and perceptual hypotheses in a wider range of species than is currently used in integrative and comparative research. Devices based on using IMU technology to measure phonotaxis represent a new line of tools that complement other types of devices designed to measure animal taxis while restricting the gross movements of subjects in space (Doherty & Pires, 1987; Hedrick, Hisada, & Mulloney, 2007; Kramer, 1976; Weber, Thorson, & Huber, 1981; Wendler, Dambach, Schmitz, & Scharstein, 1980). Most previous devices have been designed to measure walking in insects, and we are aware of only one such device previously designed for use with frogs (Márquez, Bosch, & Eekhout, 2008). We suggest that IMU devices might be implemented to solve some of the following impediments to advancing research on mate choice in frogs and perhaps some other acoustically signalling animals.

First, the low costs and high portability of small IMU devices could significantly broaden behavioural research by making the rigorous assessment of phonotaxis behaviour more accessible to a broader range of scientists. One upshot of this would be to extend phonotaxis studies to a taxonomically and geographically broader range of species, thereby facilitating comparative research. Much of what we currently know about female mate choice in frogs, for example, comes from intensive studies of a limited number of species, such as North American and European hylids (Gerhardt, 2001; Gerhardt & Huber, 2002) and a single, particularly well-studied Neotropical leptodactylid (Ryan et al., 2019). We think IMUs could be especially useful for examining mate choice preferences in the context of species recognition by female frogs in putative cryptic species complexes (sensu Littlejohn et al., 1960). To the extent to which phonotaxis studies help identify unique species, such efforts could also contribute to biodiversity assessments and conservation efforts.

Second, the ability to measure phonotaxis behaviours of subjects that are held in a small enclosure instead of freely moving around through large swaths of the sound field opens up new avenues for future research. At present, for example, there are no published studies of mate choice from awake, behaving preparations that have investigated the neural correlates of auditory perception and mating decisions in frogs. The use of IMUs like that implemented in CRoAK, when combined with electrophysiology commutators, could provide an important means to investigate the neural mechanisms of perception and decision making by enabling

simultaneous behavioural and neural recordings of females in the act of responding to sexual advertisement signals. Measuring phonotaxis in a small enclosure that precludes large movements through the sound field also extends the range of basic studies of auditory perception that become possible. For example, in studies of auditory masking or auditory grouping, movement in more traditional phonotaxis tests introduces potential confounds related to changes that occur during testing in the spatial relationships between the subject and sounds presented from one or more speakers (e.g., Bee, 2007, 2008a; Bee & Riemersma, 2008; Nityananda & Bee, 2012; Ward et al., 2013a). Such confounds could be avoided using IMUs.

Finally, the portability and customizability of IMU-based devices like CRoAK create the possibility of examining phonotaxis behaviour in more natural acoustic scenes, such as in natural breeding areas. For example, if females initiate phonotaxis movements after the offset of a call, one could potentially implement analyses akin to reverse correlation methods used in neurophysiology (e.g., Eggermont, Johannesma, & Aertsen, 1983) to correlate movements registered with an IMU device with the acoustic properties of specific calls recorded simultaneously from multiple nearby males, perhaps using a microphone array that can localize call and recover their acoustic properties (e.g., Jones, Jones, & Ratnam, 2014). Such uses might reveal a great deal about how females attend to individual males in dense aggregations.

In summary, we have shown that using a small, portable, and inexpensive device to record animal movements using an IMU can reproduce a broad range of results from earlier studies of species recognition and sexual selection in frogs. We encourage researchers to consider whether using IMUs to measure phonotaxis in frogs and other small animals might provide a preferable alternative to more traditional phonotaxis methods as a tool to answer fundamental questions in animal behaviour. Before using an IMU-based approach, researchers should also consider whether such a device can be configured to collect the data most relevant to testing their hypothesis of interest. For example, several procedures described in our manuscript (e.g., randomization of stimuli in sequences, normalization of responses across stimuli and individuals, scaling responses after a sound to those right before the sound) were incorporated to minimize the impacts of individual differences in general motivation, activity levels, and body size on the data we collected. Doing so was appropriate for this validation study but might lead to missing out on important data collection in other studies, depending on the hypothesis being tested. Understanding both how potential subjects behave inside the apparatus and what the apparatus actually measures should be important first steps in any future IMU-based study of phonotaxis behaviour.

Data Accessibility

The data sets generated during this current study are available in the Digital Repository of the University of Minnesota, https://doi.org/10.13020/fzk6-1q04.

References

Akre, K. L., & Ryan, M. J. (2010). Complexity increases working memory for mating signals. *Current Biology, 20*(6), 502-505.

Bailey, N. W., & Zuk, M. (2008). Acoustic experience shapes female mate choice in field crickets. *Proceedings of the Royal Society B: Biological Sciences, 275*(1651), 2645-2650.

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48.

Beckers, O. M., & Schul, J. (2004). Phonotaxis in *Hyla versicolor* (Anura, Hylidae): the effect of absolute call amplitude. *Journal of Comparative Physiology A, 190*, 869-876.

Bee, M. A. (2007). Sound source segregation in grey treefrogs: spatial release from masking by the sound of a chorus. *Animal Behaviour, 74*, 549-558.

Bee, M. A. (2008a). Finding a mate at a cocktail party: Spatial release from masking improves acoustic mate recognition in grey treefrogs. *Animal Behaviour, 75*, 1781-1791.

Bee, M. A. (2008b). Parallel female preferences for call duration in a diploid ancestor of an allotetraploid treefrog. *Animal Behaviour, 76*, 845-853.

Bee, M. A., & Riemersma, K. K. (2008). Does common spatial origin promote the auditory grouping of temporally separated signal elements in grey treefrogs? *Animal Behaviour, 76*, 831-843.

Bee, M. A., & Schwartz, J. J. (2009). Behavioral measures of signal recognition thresholds in frogs in the presence and absence of chorus-shaped noise. *Journal of the Acoustical Society of America, 126*(5), 2788-2801.

Bee, M. A., & Swanson, E. M. (2007). Auditory masking of anuran advertisement calls by road traffic noise. *Animal Behaviour, 74*, 1765-1776.

Bee, M. A., Vélez, A., & Forester, J. D. (2012). Sound level discrimination by gray treefrogs in the presence and absence of chorus-shaped noise. *Journal of the Acoustical Society of America, 131*(5), 4188-4195.

Bush, S. L., Gerhardt, H. C., & Schul, J. (2002). Pattern recognition and call preferences in treefrogs (Anura: Hylidae): A quantitative analysis using a no-choice paradigm. *Animal Behaviour, 63*, 7-14.

603 Diekamp, B., & Gerhardt, H. C. (1995). Selective phonotaxis to advertisement calls in the gray
604 treefrog *Hyla versicolor*: Behavioral experiments and neurophysiological correlates.
605 *Journal of Comparative Physiology A, 177*(2), 173-190.
606 Doherty, J. A., & Pires, A. (1987). A new microcomputer-based method for measuring walking
607 phonotaxis in field crickets (Gryllidae). *Journal of Experimental Biology, 130*, 524-432.
608 Draganoiu, T. I., Nagle, L., & Kreutzer, M. (2002). Directional female preference for an
609 exaggerated male trait in canary (*Serinus canaria*) song. *Proceedings of the Royal*
610 *Society B-Biological Sciences, 269*(1509), 2525-2531.
611 Eggermont, J., Johannesma, P., & Aertsen, A. (1983). Reverse-correlation methods in auditory
612 research. *Quarterly Reviews of Biophysics, 16*(3), 341-414.
613 Fellers, G. M. (1979). Aggression, territoriality, and mating behaviour in North American
614 treefrogs. *Animal Behaviour, 27*, 107-119.
615 Gall, M. D., Bee, M. A., & Baugh, A. T. (2019). The difference a day makes: breeding remodels
616 hearing, hormones and behavior in female Cope's gray treefrogs (*Hyla chrysoscelis*).
617 *Hormones and Behavior, 108*, 62-72.
618 Gerhardt, H. C. (1975). Sound pressure levels and radiation patterns of vocalizations of some
619 North American frogs and toads. *Journal of Comparative Physiology, 102*(1), 1-12.
620 Gerhardt, H. C. (1981). Mating call recognition in the green treefrog (*Hyla cinerea*): Importance
621 of two frequency bands as a function of sound pressure level. *Journal of Comparative*
622 *Physiology, 144*(1), 9-16.
623 Gerhardt, H. C. (1991). Female mate choice in treefrogs: static and dynamic acoustic criteria.
624 *Animal Behaviour, 42*, 615-635.
625 Gerhardt, H. C. (1995). Phonotaxis in female frogs and toads: Execution and design of
626 experiments. In G. M. Klump, R. J. Dooling, R. R. Fay & W. C. Stebbins (Eds.), *Methods*
627 *in Comparative Psychoacoustics* (pp. 209-220). Basel: Birkhäuser Verlag.
628 Gerhardt, H. C. (2001). Acoustic communication in two groups of closely related treefrogs.
629 *Advances in the Study of Behavior, 30*, 99-167.
630 Gerhardt, H. C. (2008). Phonotactic selectivity in two cryptic species of gray treefrogs: Effects of
631 differences in pulse rate, carrier frequency and playback level. *Journal of Experimental*
632 *Biology, 211*(16), 2609-2616.
633 Gerhardt, H. C., & Brooks, R. (2009). Experimental analysis of multivariate female choice in
634 gray treefrogs (*Hyla versicolor*): evidence for directional and stabilizing selection.
635 *Evolution, 63*(10), 2504-2512.

636 Gerhardt, H. C., & Doherty, J. A. (1988). Acoustic communication in the gray treefrog, *Hyla*
637 *versicolor:* Evolutionary and neurobiological implications. *Journal of Comparative*
638 *Physiology A, 162*(2), 261-278.
639 Gerhardt, H. C., Dyson, M. L., & Tanner, S. D. (1996). Dynamic properties of the advertisement
640 calls of gray tree frogs: patterns of variability and female choice. *Behavioral Ecology,*
641 *7*(1), 7-18.
642 Gerhardt, H. C., Dyson, M. L., Tanner, S. D., & Murphy, C. G. (1994). Female treefrogs do not
643 avoid heterospecific calls as they approach conspecific calls: Implications for
644 mechanisms of mate choice. *Animal Behaviour, 47*(6), 1323-1332.
645 Gerhardt, H. C., & Huber, F. (2002). *Acoustic communication in insects and anurans: Common*
646 *problems and diverse solutions*. Chicago: Chicago University Press.
647 Gerhardt, H. C., Tanner, S. D., Corrigan, C. M., & Walton, H. C. (2000). Female preference
648 functions based on call duration in the gray tree frog (*Hyla versicolor*). *Behavioral*
649 *Ecology, 11*(6), 663-669.
650 Gordon, N. M., & Gerhardt, H. C. (2009). Hormonal modulation of phonotaxis and
651 advertisement-call preferences in the gray treefrog (*Hyla versicolor*). *Hormones and*
652 *Behavior, 55*(1), 121-127.
653 Gupta, S., Alluri, R. K., Rose, G. J., & Bee, M. A. (2021). Neural basis of acoustic species
654 recognition in a cryptic species complex. *Journal of Experimental Biology, 224*(23),
655 jeb243405.
656 Gupta, S., Marchetto, P. M., & Bee, M. A. (2020). Customizable Recorder of Animal Kinesis
657 (CRoAK): A multi-axis instrumented enclosure for measuring animal movements.
658 *HardwareX, 8*, e00116.
659 Hanson, J. L., Rose, G. J., Leary, C. J., Graham, J. A., Alluri, R. K., & Vasquez-Opazo, G. A.
660 (2016). Species specificity of temporal processing in the auditory midbrain of gray
661 treefrogs: long-interval neurons. *Journal of Comparative Physiology A, 202*(1), 67-79.
662 Hedrick, A. V., Hisada, M., & Mulloney, B. (2007). Tama-kugel: Hardware and software for
663 measuring direction, distance, and velocity of locomotion by insects. *Journal of*
664 *Neuroscience Methods, 164*(1), 86-92.
665 Heinen-Kay, J. L., Strub, D. B., & Zuk, M. (2018). Limited flexibility in female Pacific field cricket
666 (*Teleogryllus oceanicus*) exploratory behaviors in response to perceived social
667 environment. *Ethology, 124*(9), 650-656.
668 Jones, D. L., Jones, R. L., & Ratnam, R. (2014). Calling dynamics and call synchronization in a
669 local group of unison bout callers. *Journal of Comparative Physiology A, 200*, 93-107.

Kramer, E. (1976). The orientation of walking honeybees in odour fields with small concentration gradients. *Physiological Entomology, 1*(1), 27-37.

LaBarbera, K., Nelson, P. B., & Bee, M. A. (2020). Mate choice and the 'opposite miss' to Weber's law: Proportional processing governs signal preferences in a treefrog. *Animal Behaviour, 168*, 199-209.

Li, H., Schrode, K. M., & Bee, M. A. (2022). Vocal sacs do not function in multimodal mate attraction under nocturnal illumination in Cope's grey treefrog. *Animal Behaviour, 189*, 127-146.

Littlejohn, M. J., Fouquette, M. J., & Johnson, C. (1960). Call discrimination by female frogs of the *Hyla versicolor* complex. *Copeia, 1960*, 47-49.

Márquez, R., Bosch, J., & Eekhout, X. (2008). Intensity of female preference quantified through playback setpoints: call frequency versus call rate in midwife toads. *Animal Behaviour, 75*, 159-166.

Morris, M. R., Nicoletto, P. F., & Hesselman, E. (2003). A polymorphism in female preference for a polymorphic male trait in the swordtail fish *Xiphophorus cortezi*. *Animal Behaviour, 65*, 45-52.

Morris, M. R., Rios-Cardenas, O., & Tudor, M. S. (2006). Larger swordtail females prefer asymmetrical males. *Biology Letters, 2*(1), 8-11.

Nityananda, V., & Bee, M. A. (2012). Spatial release from masking in a free-field source identification task by gray treefrogs. *Hearing Research, 285*(1-2), 86-97.

O'Loghlen, A. L., & Beecher, M. D. (1999). Mate, neighbor and stranger songs: a female song sparrow perspective. *Animal Behaviour, 58*, 13-20.

R Core Team. (2020). R: A language and environment for statistical computing. Vienna, Austria, http://www.R-project.org/: R Foundation for Statistical Computing.

Reichert, M. S., Symes, L. B., & Höbel, G. (2016). Lighting up sound preferences: cross-modal influences on the precedence effect in treefrogs. *Animal Behaviour, 119*, 151-159.

Ritke, M. E., Babb, J. G., & Ritke, M. K. (1990). Life-History of the gray treefrog (*Hyla chrysoscelis*) in western Tennessee. *Journal of Herpetology, 24*(2), 135-141.

Rose, G. J., Brenowitz, E. A., & Capranica, R. R. (1985). Species specificity and temperature dependency of temporal processing by the auditory midbrain of two species of treefrogs. *Journal of Comparative Physiology A, 157*(6), 763-769.

Rose, G. J., Hanson, J. L., Leary, C. J., Graham, J. A., Alluri, R. K., & Vasquez-Opazo, G. A. (2015). Species-specificity of temporal processing in the auditory midbrain of gray

703 treefrogs: Interval-counting neurons. *Journal of Comparative Physiology A, 201*(5), 485-
704 503.
705 Rosenthal, G. G. (2017). *Mate Choice: The Evolution of Sexual Decision Making from Microbes*
706 *to Humans*. Princeton, NJ: Princeton University Press.
707 Rosenthal, G. G., & Evans, C. S. (1998). Female preference for swords in *Xiphophorus helleri*
708 reflects a bias for large apparent size. *Proceedings of the National Academy of Sciences*
709 *of the United States of America, 95*(8), 4431-4436.
710 Ryan, M. J., Akre, K. L., Baugh, A. T., Bernal, X. E., Lea, A. M., Leslie, C., . . . Rand, A. S.
711 (2019). Nineteen years of consistently positive and strong female mate preferences
712 despite individual variation. *The American Naturalist, 194*(2), 125-134.
713 Ryan, M. J., & Keddy-Hector, A. (1992). Directional patterns of female mate choice and the role
714 of sensory biases. *American Naturalist, 139*, S4-S35.
715 Sakitt, B. (1973). Indices of discriminability. *Nature, 241*, 133–134.
716 Schul, J., & Bush, S. L. (2002). Non-parallel coevolution of sender and receiver in the acoustic
717 communication system of treefrogs. *Proceedings of the Royal Society of London Series*
718 *B-Biological Sciences, 269*(1502), 1847-1852.
719 Schwartz, J. J., Buchanan, B. W., & Gerhardt, H. C. (2001). Female mate choice in the gray
720 treefrog (*Hyla versicolor*) in three experimental environments. *Behavioral Ecology and*
721 *Sociobiology, 49*(6), 443-455.
722 Schwartz, J. J., Huth, K., & Hutchin, T. (2004). How long do females really listen? Assessment
723 time for female mate choice in the grey treefrog, *Hyla versicolor*. *Animal Behaviour, 68*,
724 533-540.
725 Snedden, W. A., & Greenfield, M. D. (1998). Females prefer leading males: relative call timing
726 and sexual selection in katydid choruses. *Animal Behaviour, 56*, 1091-1098.
727 Tanner, J. C., Ward, J. L., Shaw, R. G., & Bee, M. A. (2017). Multivariate phenotypic selection
728 on a complex sexual signal. *Evolution, 71*(7), 1742-1754.
729 Ward, J. L., Buerkle, N. P., & Bee, M. A. (2013a). Spatial release from masking improves sound
730 pattern discrimination along a biologically relevant pulse-rate continuum in gray
731 treefrogs. *Hearing Research, 306*, 63-75.
732 Ward, J. L., Love, E. K., Baugh, A. T., Gordon, N. M., Tanner, J. C., & Bee, M. A. (2015).
733 Progesterone and prostaglandin F2α induce species-typical female preferences for male
734 sexual displays in Cope's gray treefrog (*Hyla chrysoscelis*). *Physiology & Behavior, 152*,
735 280-287.

736 Ward, J. L., Love, E. K., Vélez, A., Buerkle, N. P., O'Bryan, L. R., & Bee, M. A. (2013b).
737 Multitasking males and multiplicative females: Dynamic signalling and receiver
738 preferences in Cope's grey treefrog. *Animal Behaviour, 86*(2), 231-243.
739 Weber, T., Thorson, J., & Huber, F. (1981). Auditory behavior of the cricket. I. Dynamics of
740 compensated walking and discrimination paradigms on the Kramer treadmill. *Journal of*
741 *Comparative Physiology, 141*(2), 215-232.
742 Welch, A. M., Semlitsch, R. D., & Gerhardt, H. C. (1998). Call duration as an indicator of genetic
743 quality in male gray treefrogs. *Science, 280*(5371), 1928-1930.
744 Welch, A. M., Smith, M. J., & Gerhardt, H. C. (2014). A multivariate analysis of genetic variation
745 in the advertisement call of the gray treefrog, *Hyla versicolor*. *Evolution, 68*(6), 1629-
746 1639.
747 Wendler, G., Dambach, M., Schmitz, B., & Scharstein, H. (1980). Analysis of the acoustic
748 orientation behaviour in crickets (*Gryllus campestris* L.). *Naturwissenschaften, 67*(2), 99-
749 101.
750 Wong, R. Y., So, P., & Cummings, M. E. (2011). How female size and male displays influence
751 mate preference in a swordtail. *Animal Behaviour, 82*(4), 691-697.
752
753

FIGURE LEGENDS

Figure 1. The CRoAK apparatus that was used to measure phonotaxis-related movements in response to broadcasts of acoustic stimuli. Shown here is a 3D rendering of a CAD drawing (top) and a labelled schematic (bottom). The unit consisted of a small enclosure (86 mm diameter, 60 mm height) made from plastic canvas for walls and sprouting lids for a bottom and top. The enclosure was mounted from a shelf bracket using a suspension spring, suspension bar, and thread. The bottom of the enclosure was attached to a plywood base using thread. The 9DoF IMU was mounted to the bottom of the enclosure and attached to an Arduino Nano microcontroller. The unit was placed inside a small acoustic chamber. Sounds were broadcast from a speaker located at the approximate height of the enclosure. A photograph of the unit and an animated 3D rendering showing a frog inside the enclosure are included in the Supplementary Material. Additional details about the design and construction of the unit are reported in Gupta et al., (2020).

Figure 2. Example recording from the IMU and sound detector implemented in CRoAK showing responses to conspecific (*H. chrysoscelis*) and heterospecific (*H. versicolor*) calls. The blue trace shows the normalized amplitude of the acoustic stimulus recorded with the sound detector. The red trace shows the output of the IMU in units of deg/s. Waveforms of the acoustic stimuli are shown at two scales. The areas under the curve used to compute response magnitude are illustrated by shading in the 1 s prior to and 1 s immediately following each stimulus call. Animals were free to move at any time in the apparatus. In this trace, the subject initiated an apparently random physical movement just prior to the onset of the conspecific call. A desire to discount such movements in analyses is the primary reason response magnitude was determined as the difference between the areas under the curve immediately following and prior to each stimulus call.

Figure 3. Results from Experiment 1 on species discrimination showing normalized responses to conspecific (grey bars) and heterospecific (white bars) calls as a function of signal level for female *H. chrysoscelis* tested during (a) the breeding season (N = 21) and (b) after the breeding season (N = 21). Bars depict means ± 95% confidence intervals. Responses are normalized to the sample average of median response to the standard call.

Figure 4. Results from Experiment 2 on pulse-rate preferences showing the normalized responses of females of *H. versicolor* (open circles, N = 23) and *H. chrysoscelis* (filled circles, N

= 21) in response to stimulus calls differing in pulse rate. Points depict means ± 95% confidence intervals. For each species, asterisks indicate significant differences from responses to the standard call (20 pulses/s for *H. versicolor*; 50 pulses/s for *H. chrysoscelis*). Responses for each species are normalized to the sample average of median response to its respective standard call. The means and SDs for pulse rates in local populations of *H. versicolor* (21.1 ± 2.9 pulses/s) and *H.* chrysoscelis (48.8 ± 4.4 pulses/s) are shown along the x axis (Gupta et al., 2021; Ward et al., 2013b).

Figure 5. Results from Experiment 3 on call duration preferences showing the normalized responses of females of *H. chrysoscelis* (*N* = 21) in response to stimulus calls differing in pulse number. Points depict means ± 95% confidence intervals. The fitted curve represents a logistic fit to the normalized data. Asterisks indicate significant differences from responses to the standard call (30 pulses/call). Responses are normalized to the sample average of median response to the standard call. The mean and SD for call duration is 30 ± 4 pulses/call in local populations (Ward et al., 2013b). The location of this mean and ±2.5 SD is shown along the x axis.

Figure 6. Results from Experiment 4 measuring the call amplitude preference function relating normalized response magnitudes to call amplitude for females of *H. chrysoscelis* (N = 23). Points depict means ± 95% confidence intervals. The fitted curve represents a logistic fit to the normalized data. The horizontal dashed line shows the upper bound of the 95% confidence interval around the mean normalized response measured in the no-signal sham trial. Asterisks indicate significant differences from the sham trial. Responses are normalized to the sample average of median response to the standard call.

Figure 1

Click here to access/download;Figure;Fig. 1_ANBEH-D-23-00013_1column.tif

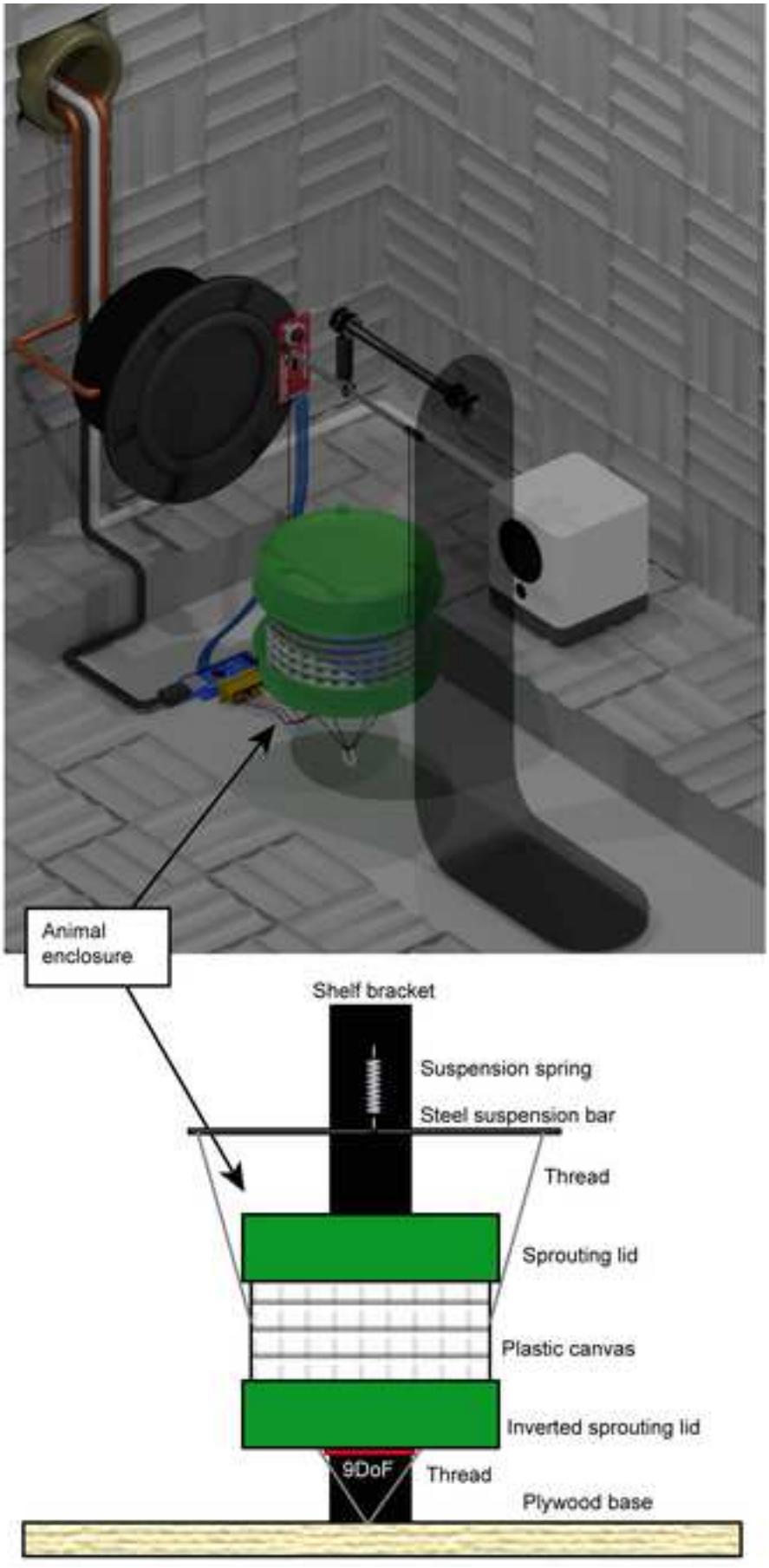

Figure 2

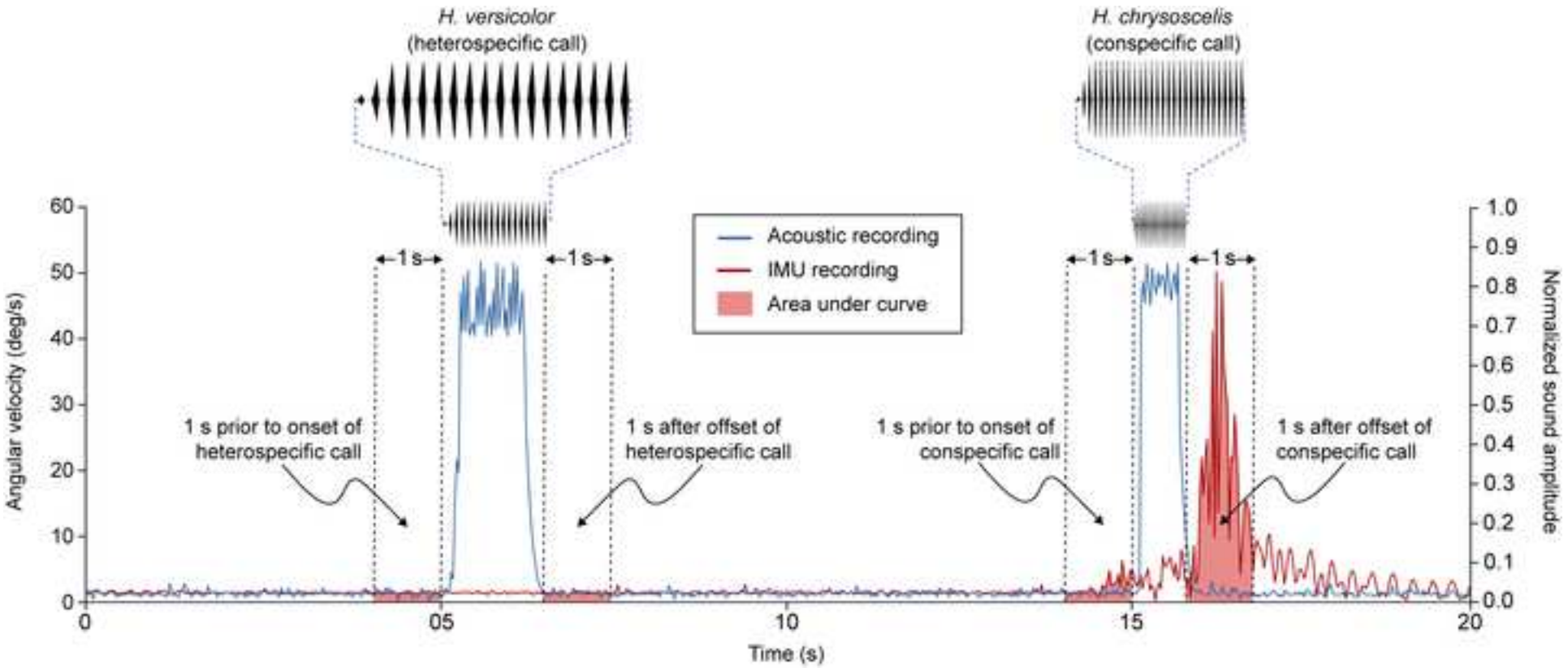

Figure 3

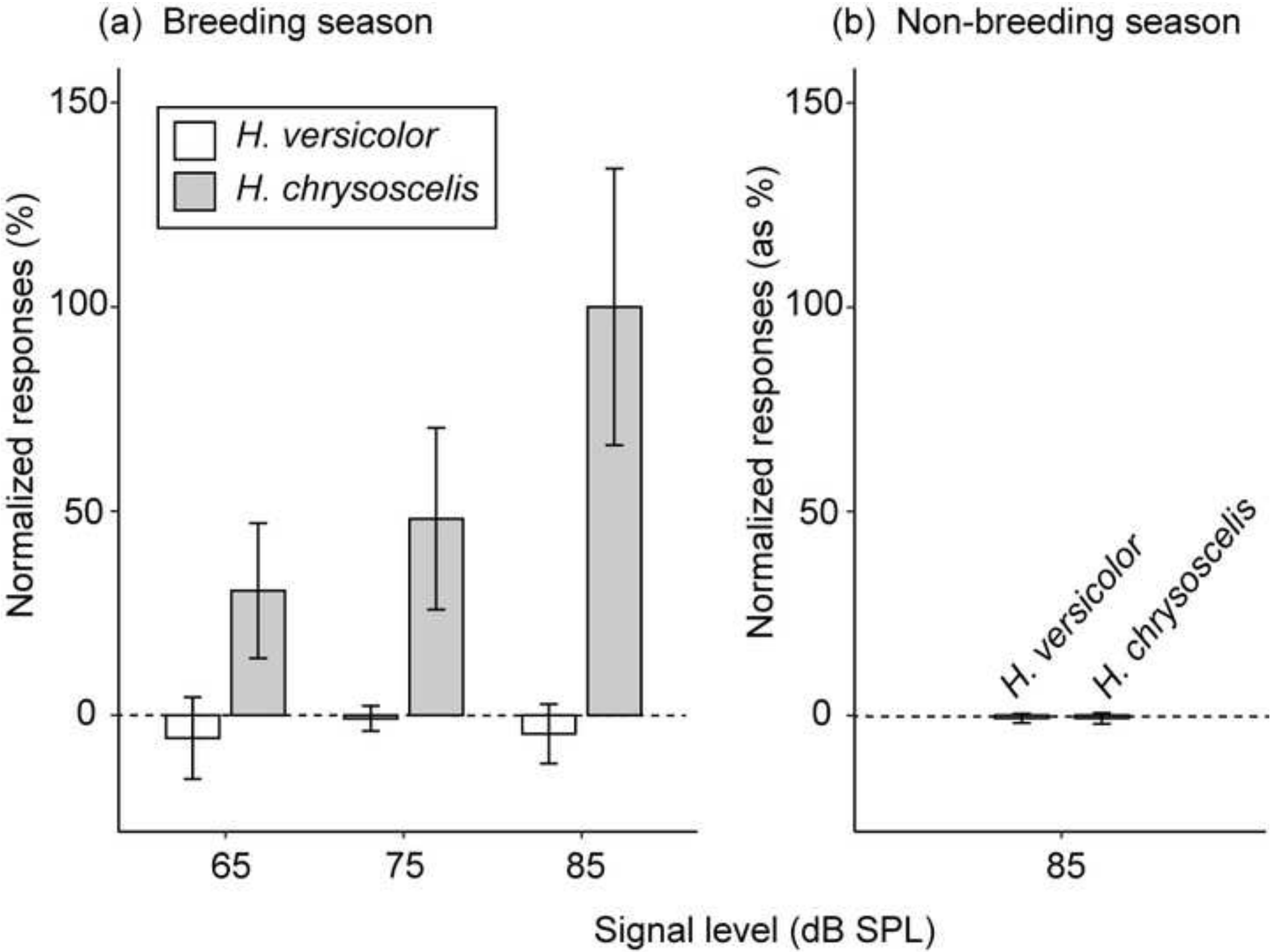

Figure 4

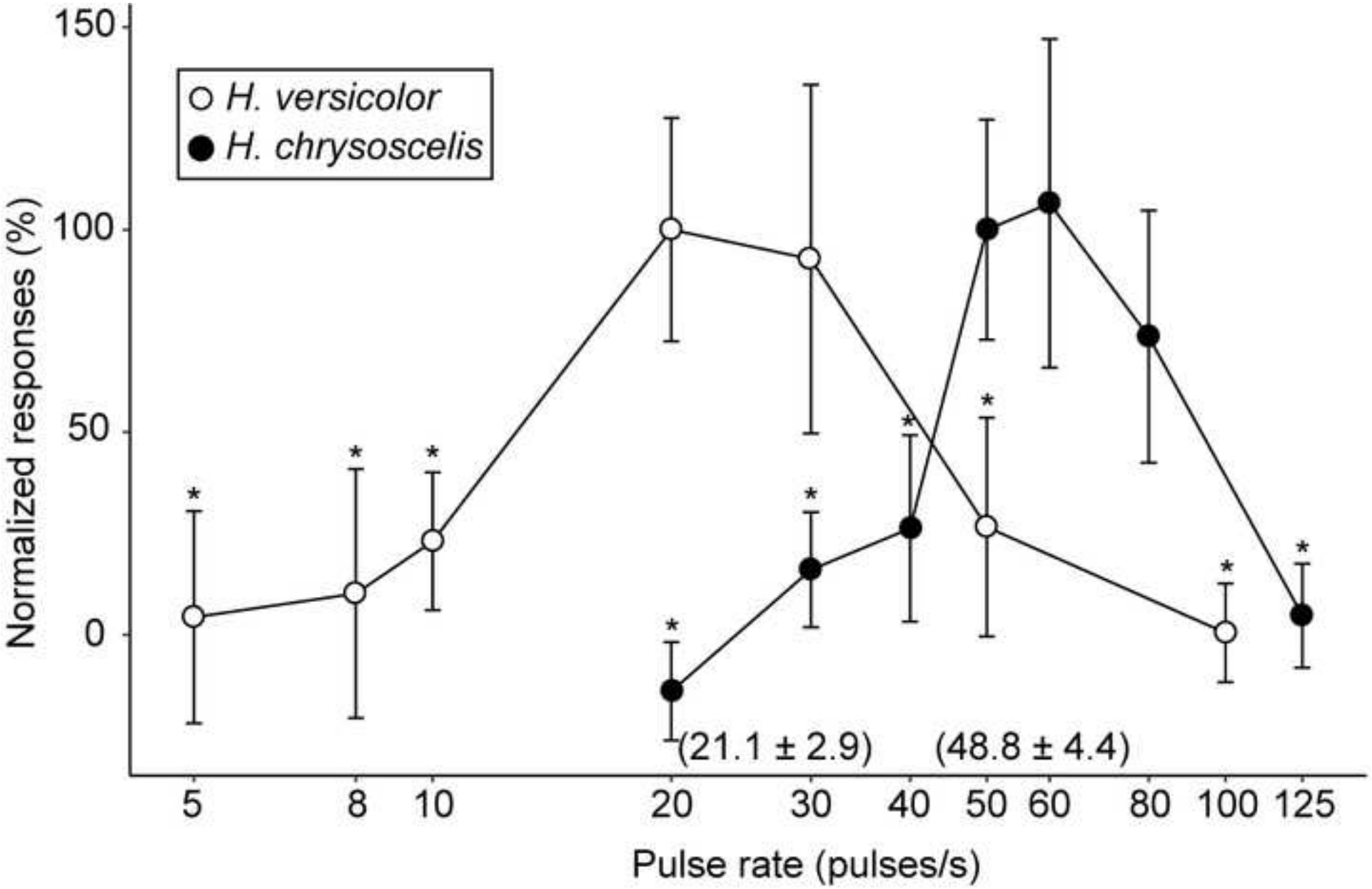

Figure 5

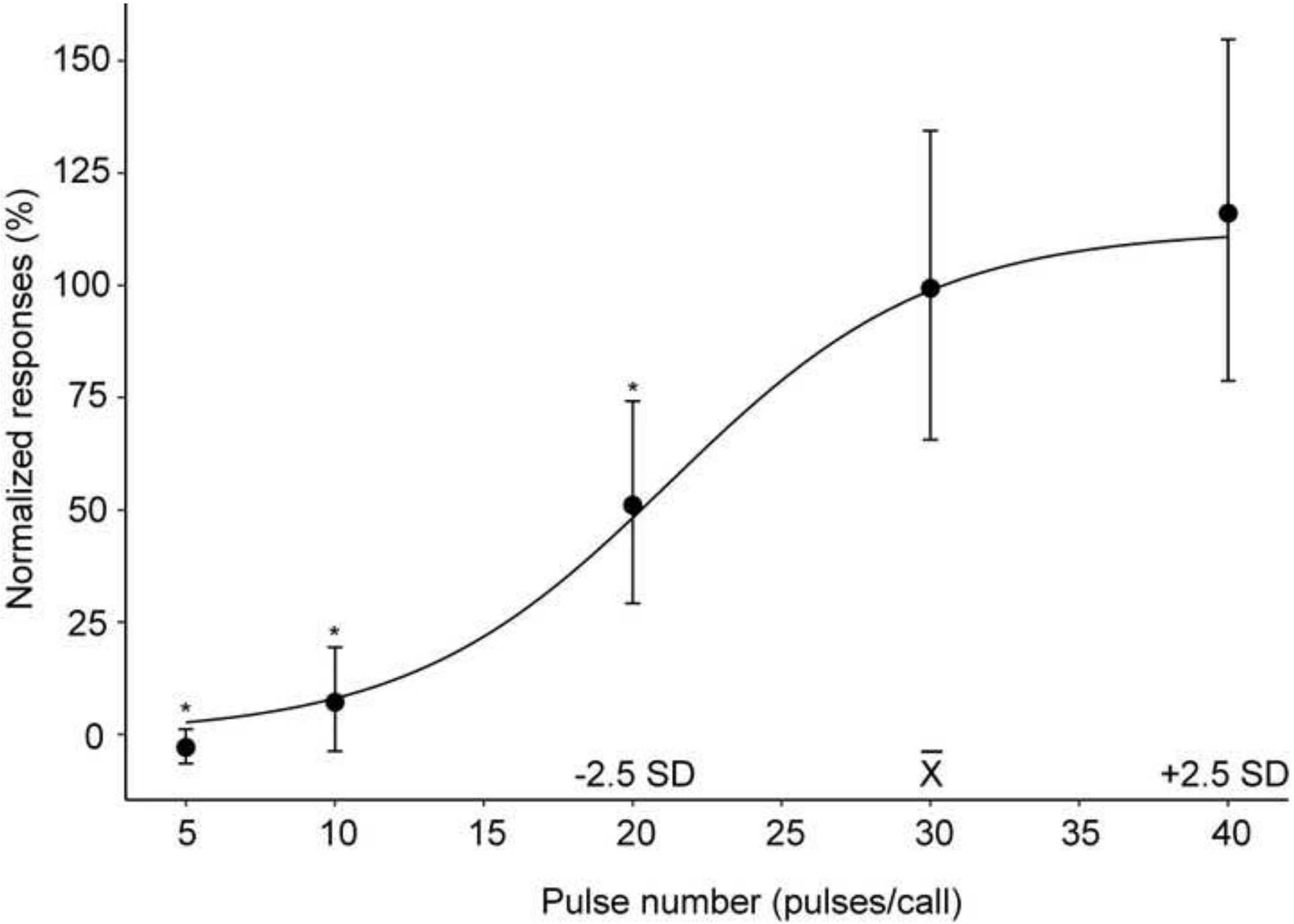

Figure 6