CrossMark

ORIGINAL PAPER

Nonlinear processing of a multicomponent communication signal by combination-sensitive neurons in the anuran inferior colliculus

Norman Lee[1,3] · Katrina M. Schrode[2,4] · Mark A. Bee[1,2]

Abstract Diverse animals communicate using multicomponent signals. How a receiver's central nervous system integrates multiple signal components remains largely unknown. We investigated how female green treefrogs (*Hyla cinerea*) integrate the multiple spectral components present in male advertisement calls. Typical calls have a bimodal spectrum consisting of formant-like low-frequency (~0.9 kHz) and high-frequency (~2.7 kHz) components that are transduced by different sensory organs in the inner ear. In behavioral experiments, only bimodal calls reliably elicited phonotaxis in no-choice tests, and they were selectively chosen over unimodal calls in two-alternative choice tests. Single neurons in the inferior colliculus of awake, passively listening subjects were classified as combination-insensitive units (27.9%) or combination-sensitive units (72.1%) based on patterns of relative responses to the same bimodal and unimodal calls. Combination-insensitive units responded similarly to the bimodal call and one or both unimodal calls. In contrast, combination-sensitive units exhibited both linear responses (i.e., linear summation) and, more commonly, nonlinear responses (e.g., facilitation, compressive summation, or suppression) to the spectral combination in the bimodal call. These results are consistent with the hypothesis that nonlinearities play potentially critical roles in spectral integration and in the neural processing of multicomponent communication signals.

Keywords Auditory midbrain · Call recognition · Complex signal · Neural integration · Nonlinear processing

Abbreviations

ANOVA	Analysis of variance
BEF	Best excitatory frequency
CV	Coefficient of variation
DFA	Discriminant function analysis
FTC	Frequency tuning curve
GABA	Gamma-aminobutyric acid
IC	Inferior colliculus
MTR	Multi-tone responder
SD	Standard deviation
SPL	Sound pressure level
STR	Single-tone responder

✉ Mark A. Bee
mbee@umn.edu

[1] Department of Ecology, Evolution, and Behavior, University of Minnesota, St. Paul, MN 55108, USA

[2] Graduate Program in Neuroscience, University of Minnesota, Minneapolis, MN 55455, USA

[3] Present Address: Biology Department, St. Olaf College, Northfield, MN 55057, USA

[4] Present Address: Department of Otolaryngology, Johns Hopkins School of Medicine, Baltimore, MD 21205, USA

Introduction

Animal communication signals commonly consist of multiple components (Partan and Marler 1999, 2005; Rowe 1999; Candolin 2003; Hebets and Papaj 2005). When receivers detect and process different signal components using different sensory systems, such as the visual and seismic components of spider courtship displays (Uetz et al. 2013; Girard et al. 2015) or the visual, acoustic, and olfactory components found in many defensive warning displays (Rowe and Halpin 2013), such signals are considered "multimodal signals" (Partan and Marler 1999, 2005; Higham and Hebets 2013). Many animal

signals, however, consist of distinct signal components that are nevertheless processed within the same sensory modality, such as the red coloration and zig-zag courtship dance of the three-spined stickleback (Wootton 1976; Milinski and Bakker 1990; Rowland 1994) or the "whine" and "chuck" of the túngara frog's advertisement call (Ryan 1985; Farris and Taylor 2016). Because they are transduced by one sensory modality, such signals are considered "unimodal multicomponent" signals (Higham and Hebets 2013), and they are used widely for communication by diverse taxa (Bradbury and Vehrencamp 2011). (Hereafter, we refer to such signals simply as "multicomponent signals.")

Over the preceding two decades, behavioral studies of communication in a diversity of animals have significantly advanced our conceptual understanding of the function and evolution of multicomponent and multimodal signals (Gerhardt 1992; Partan and Marler 1999; Rowe 1999; Hebets and Papaj 2005; Higham and Hebets 2013; Bro-Jørgensen 2010). By comparison, however, relatively few physiological studies have investigated the neural processing of multicomponent communication signals (Partan 2013; Bee and Miller 2016). One important aspect of receiving multicomponent and multimodal signals is that the processing of separate components eventually becomes integrated in the receiver's central nervous system. Consequently, *combinations* of components are processed differently than single components by themselves. Elucidating the integrative processing of component combinations by receivers' nervous systems is ultimately necessary to more fully understand multicomponent and multimodal signaling and the evolution of signal complexity.

The present study of female green treefrogs (*Hyla cinerea*, Hylidae) investigated spectral integration by neurons in the auditory midbrain in relation to the receiver's behavioral responsiveness to, and selectivity for, the multicomponent frequency spectrum of the male's advertisement call. During the breeding season, male green treefrogs (Fig. 1a) produce an advertisement call consisting of a single note of about 100–200 ms in duration (Fig. 1c). Actively calling males emit an advertisement call approximately every 1–2 s from stationary positions on emergent vegetation, and females select their mate by approaching calling males (Fig. 1b; Gerhardt 2001). The frequency spectrum of the call is considered "bimodal" (in contrast to "unimodal") because it has two prominent, formant-like spectral components (Fig. 1c; Gerhardt 1974, 1976, 1981b; Oldham and Gerhardt 1975). (Hereafter, we use "bimodal" and "unimodal" to describe the number of modes in a sound's frequency spectrum.) The low-frequency component is centered between 0.64 and 1.34 kHz (typically close to 0.9 kHz) and the high-frequency

component is centered between 2.2 and 3.6 kHz (typically close to 2.7–3.0 kHz) (Oldham and Gerhardt 1975; Gerhardt 2001). Previous work suggests the two spectral components may have different primary functions. The low-frequency component appears to be relatively more important for source localization and long-distance attraction, whereas the high-frequency component may be more critical in closer-range mate choice decisions (Gerhardt 1976, 1981b; Rheinlaender et al. 1979; Klump et al. 2004).

In two behavioral experiments, we evaluated phonotaxis responses (Gerhardt 1995) to synthetic advertisement calls designed to have either a bimodal spectrum with both components or a unimodal spectrum consisting of just the low-frequency or the high-frequency spectral component. In Experiment 1, we used two-alternative choice tests to investigate behavioral selectivity for bimodal and unimodal calls. Experiment 2 consisted of single-stimulus, no-choice tests that investigated responsiveness to bimodal and unimodal calls by determining the extent to which they were recognized as behaviorally salient signals capable of eliciting phonotaxis. Both experiments were designed to build on previous studies investigating the role of spectral cues in call recognition and discrimination in green treefrogs (Gerhardt 1974, 1976, 1981b, 1986). In Experiment 3, we used extracellular, single-unit recordings to investigate the responses to bimodal and unimodal calls of neurons in the torus semicircularis, the evolutionary homolog of the mammalian inferior colliculus (IC) (Nieuwenhuys et al. 1998). This experiment was designed to characterize the extent to which neurons in the frog IC exhibit combination-sensitive responses (both linear and nonlinear) to the combination of spectral components in bimodal calls compared to unimodal calls. Combination sensitivity is a common feature of vertebrate auditory processing that plays a critical role in spectral integration (Margoliash and Fortune 1992; Mittmann and Wenstrup 1995; Kadia and Wang 2003; Akimov et al. 2017). Combination-sensitive responses can also be found in the frog's auditory thalamus (Fuzessery and Feng 1983; Megela 1983; Mudry and Capranica 1987a, b), suggesting the thalamus is an important region in the anuran brain where integration of separate spectral components occurs. Previous studies of anurans, however, have produced conflicting results as to whether similar integrative processes occur at the level of the IC (Mudry et al. 1977; Fuzessery and Feng 1982). Moreover, these studies have been conducted using a species (northern leopard frogs, *Rana pipiens*) for which no data on behavioral responses to spectral combinations exist. Our study, therefore, aimed to relate a receiver's behavior in response to a relatively simple yet multicomponent acoustic signal to processing by neurons in the same species' auditory midbrain.
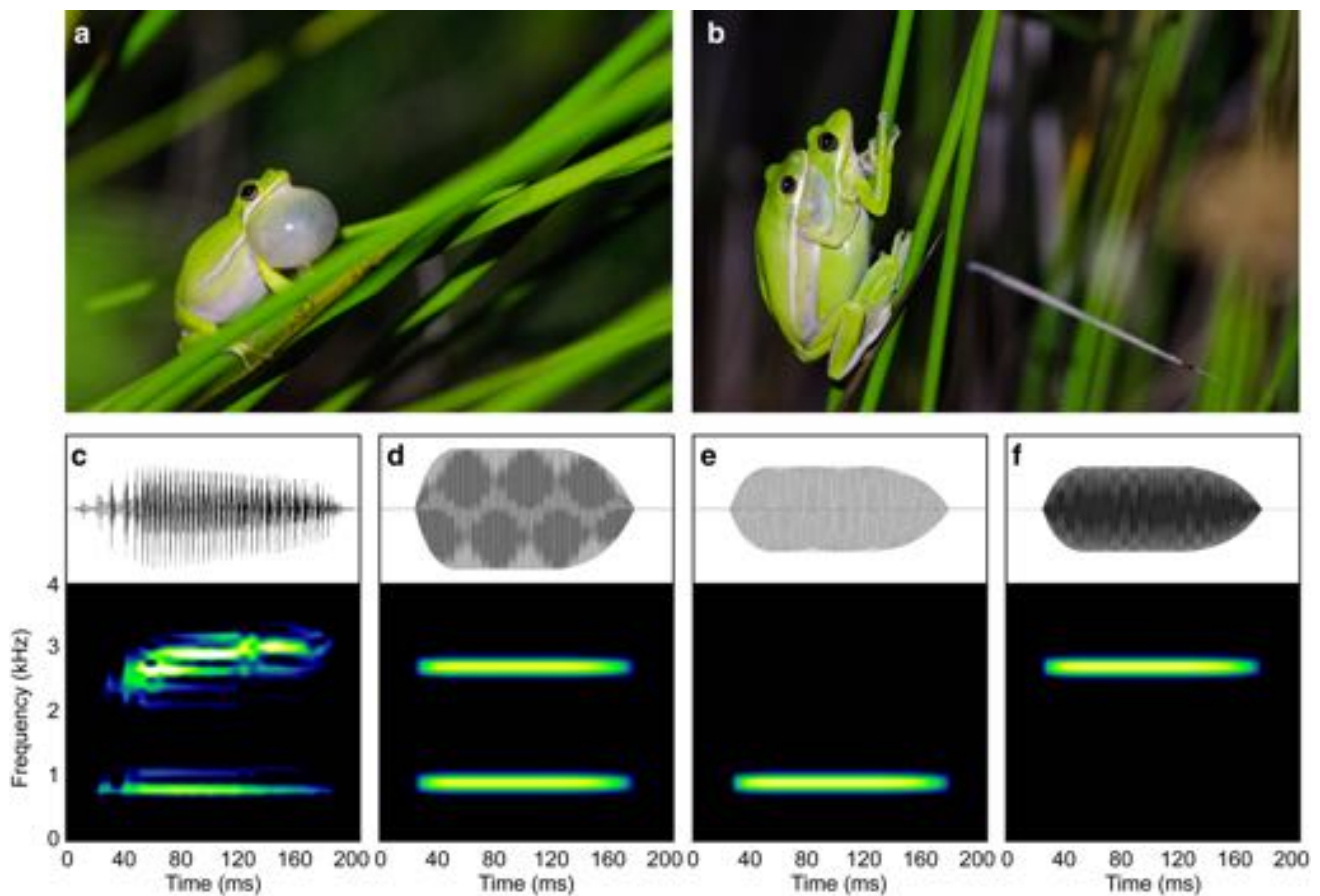
Fig. 1 Green treefrogs and natural and synthetic advertisement calls. **a** A calling male green treefrog produces as advertisement call. **b** A pair of green treefrogs in amplexus (male, *left*; female, *right*). **c–f** Oscillograms (*top*) and spectrograms (*bottom*) of **c** a natural *Hyla cinerea* advertisement call recorded from a male in our study popula-tion, **d** the synthetic bimodal call consisting of spectral components at 0.9 and 2.7 kHz, **e** the synthetic 0.9-kHz unimodal call, and **f** the synthetic 2.7-kHz unimodal call. The unimodal calls depicted in **e** and **f** have the same amplitude as the corresponding spectral component in the bimodal call depicted in **d**

Methods

Subjects

We used wild-caught females as subjects in all experiments. Collections were made at night, in active choruses, between 25 April and 15 May, in 2013–2016. Subjects were collected in amplexus (Fig. 1b) in wetlands managed by the East Texas Conservation Center in Jasper County, Texas, USA (30°56′46.15″N, 94°7′51.46″W). Collected subjects were maintained at 2–4 °C to prevent egg laying and returned to the laboratory and tested in a behavioral experiment within 48–72 h. After behavioral testing, subjects were either returned to their collection site within 2 days and released unharmed, or they were maintained in the laboratory on a 12:12 light:dark cycle, fed vitamin-dusted crickets, and given ad libitum access to water until used as a subject in a neurophysiological experiment. Collections were made under scientific permit #SPR-0410-054 issued by the Texas Parks & Wildlife Department.

Behavioral methods

Protocols for our behavioral experiments followed procedures outlined in more detail elsewhere (e.g., Ward et al. 2013a, b; Lee et al. 2017). Phonotaxis tests were conducted in an acoustically transparent, circular arena (2-m diameter) located inside a temperature-controlled, hemi-anechoic sound chamber (Industrial Acoustics Corporation, Bronx, New York, USA; inside dimensions: 300 cm × 280 cm × 216 cm). Tests were conducted at 24 ± 1 °C, which approximates a typical night-time temperature during the breeding season at the field site where animals were collected. Acoustic stimuli were output from a PC (Dell Optiplex 960, Dell Corporation, Round Rock, Texas, USA) through an external sound card (Firewire 410, M-Audio, Cumberland, Rhode Island, USA or MOTU16A,

Mark of the Unicorn, Cambridge, Massachusetts, USA), amplified (HTD DMA-1275, Home Theater Direct Inc., Plano. Texas, USA, or Crown XLS1000, Crown Audio, Elkhart, Indiana, USA), and presented through either one or two speakers (Mod1, Orb Audio LLC, New York, New York, USA). Speakers were placed on the carpeted floor of the sound chamber outside the arena walls and directed toward a subject release cage 1 m away in the center of the arena. At the start of a test, a single subject was placed in the release cage and given a 1-min acclimatization period after which stimulus playback commenced. After approximately 30 s of stimulus playback, the subject was released remotely from outside the chamber. Tests were conducted under infrared illumination and could be viewed on a monitor located outside the chamber using an infrared-sensitive camera mounted above the center of the test arena from the chamber ceiling. Sound pressure levels (dB SPL re 20 μPa) were calibrated at a distance of 1 m from the sound source by placing the microphone (Brüel & Kjær Type 4950, Brüel & Kjær, Nærum, Denmark) of a sound level meter (Brüel & Kjær Type 2250) at the approximate position of a subject's head at the start of a phonotaxis trial.

Both experiments used the same three acoustic stimuli (44.1 kHz, 16 bit) (Fig. 1d–f), which consisted of synthetic calls generated in Matlab (v2014a, MathWorks, Natick, Massachusetts, USA). These stimuli were designed to have species-typical spectral and temporal properties based on analyses of the advertisement calls (Fig. 1c) produced by male green treefrogs in previous studies (Gerhardt 2001), which were also similar to values recorded in our study population (unpublished data). A bimodal call (Fig. 1d) was generated by adding two phase-locked sinusoids (starting phase 0°) with equal amplitudes and frequencies of 0.9 and 2.7 kHz. Our bimodal call lacked the more prominent 300-Hz periodicity generated in synthetic calls with spectral components at 0.9, 2.7, and 3.0 kHz, which have been used in some previous studies (e.g., Gerhardt 1981b; Klump et al. 2004). We also generated two unimodal calls using either the 0.9-kHz (Fig. 1e) or the 2.7-kHz (Fig. 1f) sinusoid alone. All stimuli were 150 ms in duration and were shaped with 25-ms on-ramps that reached 50% amplitude in 25% of the rise time, and 50-ms off-ramps that reached 50% amplitude at 75% of the fall time. Synthetic stimuli were presented at naturalistic amplitudes (Gerhardt 1975) and were repeated at rates of 40 calls/min (1.5-s call period) to simulate actively calling males.

Experiment 1 consisted of four two-alternative choice tests in which subjects were allowed to choose between the bimodal call and one of the unimodal calls. The purpose of this experiment was to investigate receiver selectivity for bimodal versus unimodal calls. The order of the four two-choice tests was randomized for each subject. For the bimodal alternative, each spectral component was independently calibrated to a sound pressure level of 82 dB SPL so that the resulting overall level of the bimodal call was 85 dB SPL. In two of the four choice tests, the unimodal alternative was calibrated to a level of 82 dB SPL so that its overall level matched the level of the corresponding spectral component in the bimodal call. In the remaining pair of choice tests, each unimodal alternative was calibrated to 85 dB SPL so that its overall level matched the overall level of the bimodal call. This second pair of tests allowed us to control for the 3-dB difference in overall level between the bimodal and unimodal calls in the first pair of tests. The two alternative calls in a particular test alternated in time with equal periods of silence between them. The two speakers used in choice tests were placed 90° apart around the perimeter of the circular test arena. Using a factorial design, we counterbalanced across subjects whether the bimodal call was presented from the left or the right speaker and whether it was the first or the second stimulus presented in the test. We scored a choice when the subject entered a 10-cm wide response zone in front of one of the two speakers within 10 min. We analyzed the outcome of each two-alternative choice test ($n = 32$/test) using two-tailed binomial tests of the hypothesis that the two alternatives were chosen in equal proportions of 0.5.

In Experiment 2, we conducted six no-choice tests in which only a single stimulus was presented. The purpose of this experiment was to compare receiver responsiveness to bimodal and unimodal calls. In the first and last tests, we presented the bimodal call calibrated to 85 dB SPL. In the intervening four tests, we used a factorial design to present each of the two unimodal calls (0.9 or 2.7 kHz) at each of two levels (82 or 85 dB SPL) in an order that was randomized for each subject. We scored a response and noted the latency to respond when a subject entered the response zone in front of the active speaker. Subjects ($n = 32$) were given up to 5 min to respond. (Results from the first experiment indicated that subjects typically responded in less than 5 min.) We converted each individual's response latencies to phonotaxis scores (Bush et al. 2002) by dividing that individual's average latency to respond to the bimodal calls in the first and last tests by their latency to respond in each of the six tests. Relatively higher phonotaxis scores correspond to relatively faster responses to presumably more attractive stimuli. Subjects that failed to respond to one of the unimodal calls were assigned a phonotaxis score of 0 for that test. We compared phonotaxis scores across the six no-choice tests using repeated-measures analysis of variance (ANOVA) and planned contrasts.

Neurophysiological methods

Surgery and animal preparation

Approximately 24 h prior to conducting an electrophysiological recording in Experiment 3, subjects were anesthetized using MS-222 (250 mg/L, Western Chemical Inc., Ferndale, Washington, USA) and a craniotomy was performed to expose the dura covering the optic tectum. Following loss of the leg withdrawal reflex in response to toe pinch, the animal was draped in moist surgical gauze to facilitate cutaneous respiration, and a combination of local anesthetic (2.5% lidocaine HCl) and antiseptic (0.13% benzalkonium Cl, WellSpring Pharmaceutical Corp., Sarasota, Florida, USA) was applied topically to the incision area. A small incision in the skin of the head was made to expose the dorsal skull. A drill (OmniDrill35, World Precision Instruments, Sarasota, Florida, USA) was used to thin the skull above the optic tectum until the remaining bone was thin enough to remove using forceps and fine surgical scissors. After performing the craniotomy, the hole was filled with Gelfoam® (Pfizer Inc., New York, New York, USA), the skin was put back into place, and the animal was allowed to recover overnight from the effects of anesthesia. On the day of electrophysiological recordings, the animal was immobilized with an intramuscular injection of succinylcholine chloride (5 µg/g, Sigma-Aldrich, St. Louis, Missouri, USA) and draped in moist gauze. Additional local anesthetic/antiseptic was applied to the skin before exposing the opening of the skull and cutting through the dura. The animal was then positioned in a natural posture on a small platform mounted on a vibration isolation table (TMC 63-533; Technical Manufacturing Corporation, Peabody, Massachusetts. USA) inside a second temperature-controlled, hemi-anechoic sound chamber (Industrial Acoustics Corporation; inside dimensions: 220 cm × 280 cm × 216 cm). The temperature inside the sound chamber was maintained at 24 ± 1 °C during recordings.

Electrophysiology

We made extracellular, single-unit recordings using glass microelectrodes fashioned from borosilicate glass capillary tubes (1.0 mm OD, 0.75 mm ID; World Precision Instruments) with a Flaming/Brown type micropipette puller (P-97, Sutter Instrument Company, Novato, California, USA). Electrodes were pulled to have resistances ranging between 1.0 and 2.0 MΩ when filled with 2 M KCl or 2 M NaCl. Using external landmarks and stereotaxis, we lowered electrodes through the optic tectum into the IC using a motorized microdrive (MP-225, Sutter Instrument

Company). We searched for well-isolated single units using the bimodal and unimodal calls described earlier. By design, our use of the bimodal and unimodal calls as search stimuli biased our sample of recorded IC units to include those that responded to the acoustic stimuli used in our behavioral experiments. Once a single unit was isolated, we initiated the experimental acoustic stimulation protocol described in the next section. We recorded data from 140 single units in 63 subjects.

At the completion of electrophysiological recordings, subjects either were euthanized in MS-222 or underwent transcardial perfusion to fix brain tissue for a subsequent histological verification of electrode position. We marked electrode tip positions using horseradish peroxidase (5% in 0.1 M phosphate buffer, pH 7.8; Sigma-Aldrich), which we iontophoresed into the brain (300 nA pulsed positive current for 5 min; 10-s pulse period, 5-s pulse duration) through the recording electrode immediately following the recording of single-unit responses to our experimental stimuli. For perfusions, we deeply anesthetized the animal in MS-222, performed a thoracotomy to expose the heart, and perfused 0.09% NaCl (10 mL) followed by fixative (30 mL 2.5% glutaraldehyde in 0.1 M phosphate buffer, pH 7.4) and then a 5% sucrose solution in fixative through a needle inserted into the ventricle. The brain was then removed, placed in a solution of 10% sucrose in fixative for 1 h, and then transferred to 30% sucrose in fixative and stored for 24 h at 4 °C. After 24 h, the brain was then cryosectioned (50 µm slices), reacted with 3,3′-diaminobenzidine (Sigma-Aldrich) to reveal horseradish peroxidase activity, mounted on Gold Seal UltraStick slides (Thermo Fisher Scientific, Waltham, Massachusetts, USA) and finally stained with cresyl violet. Histological examinations indicated that our recordings were restricted to the principal and magnocellular nuclei of the anuran IC.

Acoustic stimulation and data acquisition

We recorded neural responses to free-field presentations of the same synthetic bimodal and unimodal calls used in our behavioral experiments. As in the behavioral experiments, the 0.9-kHz and 2.7-kHz spectral components composing the bimodal call were independently calibrated to 82 dB SPL at the position of a subject's head in the recording apparatus, so that its overall level was 85 dB SPL. We calibrated each unimodal call to have the same amplitude (82 dB SPL) as the corresponding component in the bimodal call. Sound level calibrations were made using the same equipment described earlier. Stimulus presentation and data acquisition were controlled by custom-written software (StimProg v6, NL) running in Matlab. Acoustic signals (44.1 kHz, 16 bit) were output through a

digital-to-analog data acquisition (DAQ) device (NI USB 6259, National Instruments, Austin, Texas, USA), amplified (Crown XLS1000), and broadcast through a speaker (A'Diva, Gallo Acoustics, San Antonio, Texas, USA) positioned 75 cm to the left or right of the subject at the level of the tympanum. Neural responses were recorded from the side of the IC contralateral to the speaker. The biological signal was amplified with a Dagan 2400A extracellular amplifier (Dagan Corporation, Minneapolis, Minnesota, USA) and digitized at a sampling rate of 44.1 or 48.1 kHz using the same DAQ device used to output acoustic stimuli. Neural responses were visualized on an oscilloscope and monitored acoustically during recordings.

Upon isolating a single unit, we generated an excitatory frequency tuning curve (FTC) and recorded responses to the bimodal and unimodal calls. FTCs were generated by presenting tone pips (50-ms duration, 10-ms rise and fall times) that varied in frequency (19 values between 125 Hz and 8 kHz, 1/3-octave steps) and sound pressure level. We used an adaptive tracking procedure to determine the response threshold at each frequency. Starting at an initial sound pressure level of 65 dB SPL, the sound level was increased or decreased by 10-dB steps depending on whether the unit responded with at least one action potential in response to two out of four consecutive tone presentations (1-s tone period). For every reversal in response (i.e., from response to no response and vice versa), the step size was halved until a 2.5-dB step size was reached. The lowest sound level to evoke responses was considered the upper bound of the threshold estimate and the next lowest level was used as the lower bound. Threshold was computed on a linear scale as the average of the upper and lower bounds.[1] A custom-written, online window discriminator counted spikes exceeding a threshold voltage during each tone presentation. We used cubic splines to find the best-fit function to describe the FTC ($0.875 \leq R^2 \leq 0.999$; mean $R^2 = 0.98$), from which we determined the unit's best excitatory frequency (BEF), threshold, bandwidth at 10 dB above threshold, and Q_{10dB} value.

We recorded a unit's response to a minimum of 5 repetitions of each synthetic call (1.5-s call period); responses to 10 or more presentations of each call were available for most cells. The order of call presentation was randomized for each subject, but we obtained all responses to one stimulus before beginning presentations of the next one. Whether we first generated the FTC or presented call stimuli varied across subjects, resulting in a smaller number of units ($n = 87$) for which both responses to calls and

FTCs were obtained due to our inability to hold units long enough to obtain both for all neurons.

For a small subset of units, we were also able to measure excitatory and inhibitory response maps. Excitatory response maps were generated by presenting tones at factorial combinations of frequency (19 values between 125 Hz and 8 kHz, 1/3-octave steps) and sound pressure level (15 levels between 20 dB and 90 dB, 5-dB steps). Inhibitory response maps were generated by presenting an excitatory tone at the unit's BEF and 10 dB above threshold while simultaneously presenting potential inhibitory tones using the same factorial combinations of 19 frequencies and 15 levels. Spike counts from these recordings were averaged over 3–5 tone repetitions and interpolated across frequency and sound pressure level to generate response maps.

Data analysis

An offline, custom-written window discriminator was implemented in Matlab to register the time of occurrence of spikes exceeding a specified threshold voltage within a 250-ms window centered over each presentation of the 150-ms bimodal and unimodal calls. Analysis windows were offset by the travel time necessary for sound to reach the tympanum. A unit's spontaneous firing rate (spikes/s) was estimated by counting spikes over the 50-ms silent period preceding the onset of each presentation of a bimodal or unimodal call. From acoustically driven responses to the bimodal and unimodal calls, we determined the mean number of spikes/stimulus, the mean first-spike latency, and the mean and variance of the inter-spike interval during each stimulus presentation. Not all measures could be determined for all 140 units because some units did not respond to all three stimuli. We used repeated-measures ANOVA to compare these measures of spiking activity in response to the bimodal and unimodal calls.

Following previous studies (e.g., Fuzessery and Feng 1982), we aimed to classify units based on their responses to bimodal and unimodal calls. To this end, we devised a hierarchical classification scheme based on the relative magnitudes of responses to bimodal and unimodal calls. For each unit, we expressed its responses (mean spikes/stimulus) to the 0.9-kHz and 2.7-kHz unimodal calls as percentages of its response to the bimodal call (which, by definition, was 100%). We then implemented a series of dichotomous decision rules that allowed us to classify units at each of three hierarchical levels, the highest of which was combination-sensitive versus combination-insensitive. Our decision rules were derived from visual inspections of peristimulus time histograms. All units classified as combination-sensitive according to this classification scheme were further

---

[1] Averaging upper and lower bounds differing by 2.5 dB on a linear scale yields threshold estimates that are 0.1775 dB higher than those averaged on a dB scale.

compared by computing the index of interaction (*I*) according to the following formula:

$$I = \frac{(R_b - (R_{0.9} + R_{2.7}))}{(R_b + R_{0.9} + R_{2.7})},$$

where R_b, $R_{0.9}$, and $R_{2.7}$ are the neuron's responses (mean spikes/stimulus) to the bimodal call, the 0.9-kHz unimodal call, and the 2.7-kHz unimodal call, respectively (e.g., Yan and Suga 1996; Leroy and Wenstrup 2000). The index of interaction represents a standardized measure of the extent to which a unit's response to the combination of 0.9 and 2.7 kHz in the bimodal call exceeded or fell below expected values based on the sum of responses to these two frequencies presented separately in unimodal calls. We compared *I* values across subclasses of combination-sensitive units using a one-way ANOVA. For each subclass, we also used a two-tailed, one-sample *t* test to evaluate whether the mean value of *I* was significantly different from a null expectation of 0.0, which corresponds to perfect linear summation of responses to the 0.9-kHz and 2.7-kHz spectral components.

We used discriminant function analysis (DFA) to evaluate the extent to which measures of spiking activity and frequency tuning could be used to assign units to the same classifications derived using the dichotomous decision rules of our hierarchical classification scheme. The purpose of these analyses was to assess whether various neural response measures that were independent of the relative comparisons of standardized responses made by our decision rules could nevertheless classify units to the same categories as the decision rules. A series of 12 DFAs was conducted to examine classification success using four sets of response measures for each of three different levels of the classification hierarchy. Classification success was computed using a leave-one-out cross-validation procedure. Three DFAs for each level of the hierarchy included input variables related to spike rate (4 variables: spontaneous rate [spikes/s] and responses to bimodal and unimodal calls [spikes/stimulus]), spike timing (9 variables: mean first-spike latency and mean and variance of inter-spike interval in responses to bimodal and unimodal calls), or frequency tuning (4 variables: BEF, threshold, 10-dB bandwidth, and Q_{10dB}). The fourth DFA for each level of the hierarchy included all 17 of these measures as input variables. We followed Titus et al. (1984) to derive the classification success expected by chance, after correcting for differences in group sizes, and we report Cohen's kappa as a statistical measure of chance-corrected classification success.
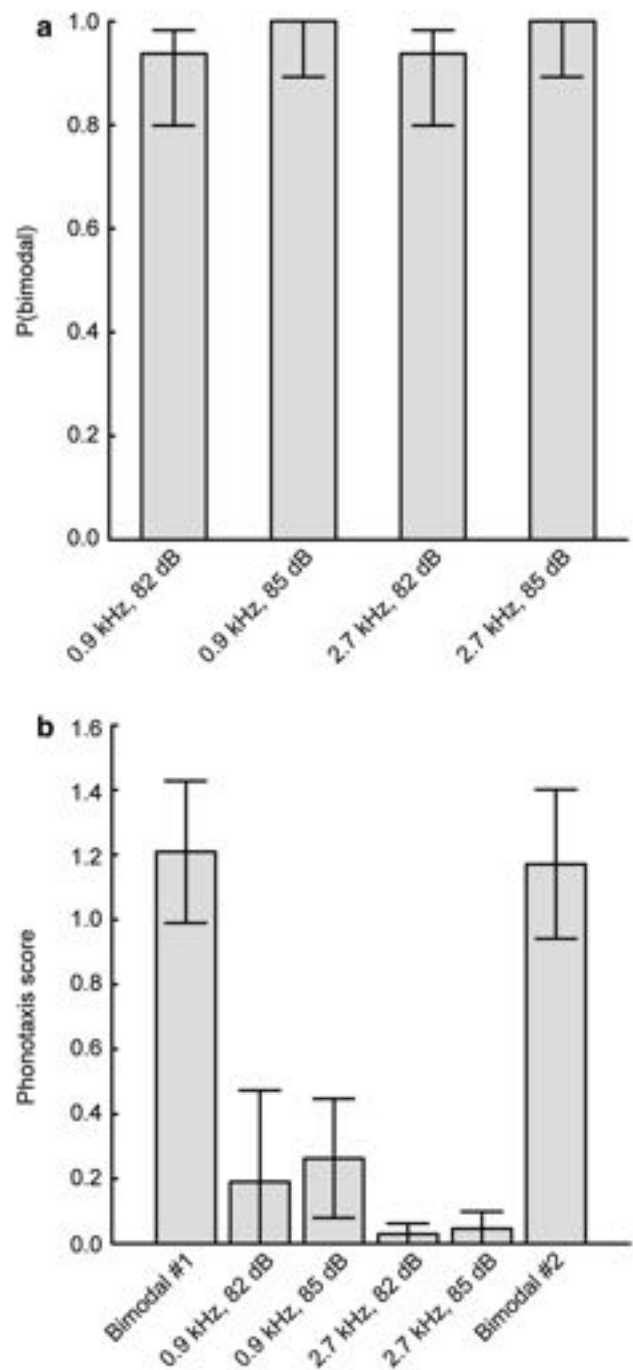

Fig. 2 Behavioral responses to bimodal and unimodal calls. **a** P(bimodal) is the proportion (±95% exact binomial confidence intervals) of females that chose the bimodal call over the unimodal call alternative indicated on the *x*-axis in two-alternative choice tests in Experiment 1. **b** Mean (±95% confidence interval) phonotaxis scores in response to bimodal and unimodal calls presented in single-stimulus, no-choice tests in Experiment 2. See text for details on computing phonotaxis scores
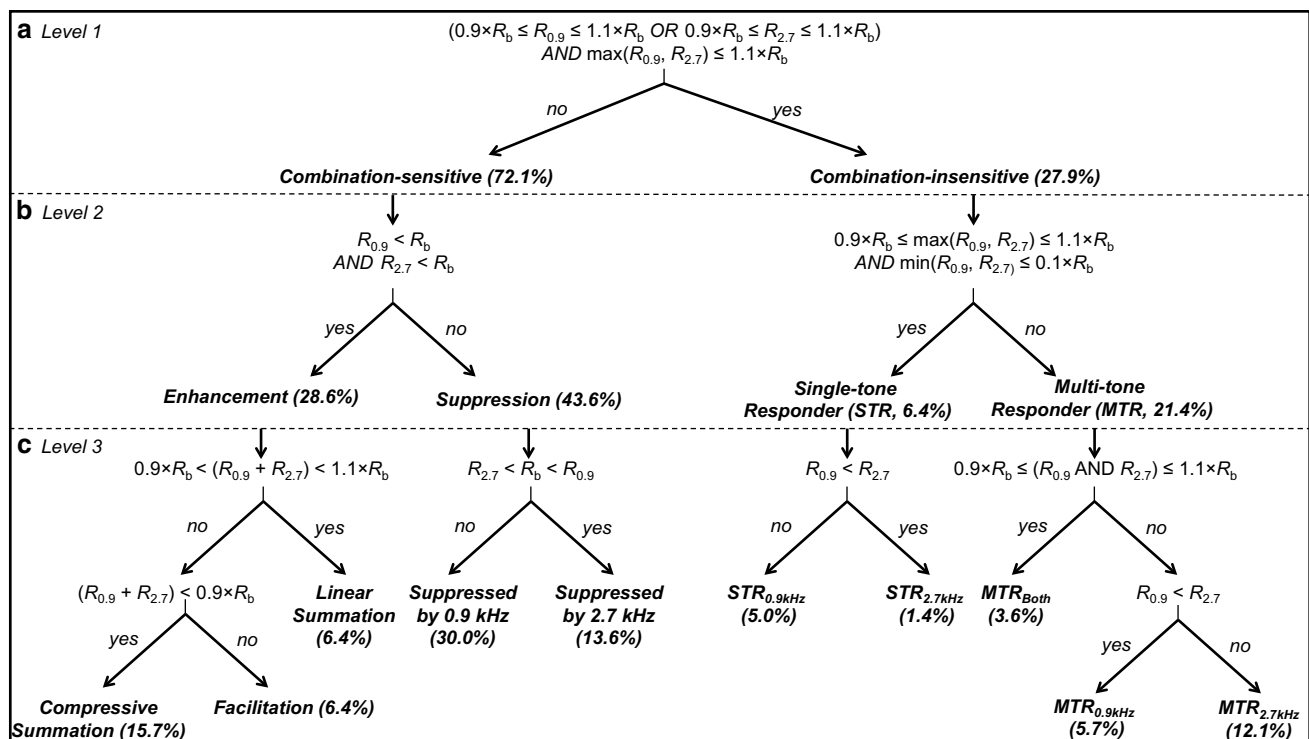
Fig. 3 Hierarchical classification scheme used to classify single units based on dichotomous decision rules. The rules combine Boolean expressions and operators to assign units to one of 10 unique classifications based on the relative response (R, in spikes/stimulus expressed as a percentage of the response to the bimodal call) to the bimodal call ($R_b = 100\%$), the 0.9-kHz unimodal call ($R_{0.9}$), and the 2.7-kHz unimodal call ($R_{2.7}$). **a** The highest level of the classification scheme classified units as combination-sensitive or combination-insensitive units. **b** At the second level of the classification scheme, combination-sensitive units were subdivided into units that displayed enhancement or suppression of their responses to the bimodal call relative to unimodal calls, and combination-insensitive units were classified as single-tone responders (STRs) or multi-tone respond-ers (MTRs) based on whether they responded to one or both of the unimodal calls. **c** At the lowest level of the classification scheme, combination-sensitive units displaying enhancement were further subdivided into those exhibiting facilitation, linear summation, or compressive summation, and those combination-sensitive units displaying suppression were subdivided based on whether they were suppressed by the addition of the 0.9-kHz or 2.7-kHz spectral component to create the bimodal call. Combination-insensitive STRs and MTRs were further subdivided based on whether they responded only (STRs) or most strongly (MTRs) to the 0.9-kHz or 2.7-kHz spectral component, or both. Percentages of all units ($n = 140$) classified as the indicated type of unit are included in parentheses

Results

Behavioral responses to unimodal and bimodal calls

In Experiment 1, subjects overwhelmingly chose bimodal calls over unimodal calls (Fig. 2a). When the sound level of the 0.9-kHz unimodal call (82 dB SPL) matched that of the corresponding component in the bimodal call, 30 of 32 (93.8%) subjects chose the bimodal call ($P < 0.001$). Likewise, 30 of 32 (93.8%; $P < 0.001$) subjects chose the bimodal call when the level of the 2.7-kHz unimodal call (82 dB SPL) matched that of the corresponding component in the bimodal call. When the overall levels of the unimodal and bimodal alternatives were equalized to 85 dB SPL, subjects unanimously (32 of 32) chose the bimodal call in both tests (0.9-kHz alternative: $P < 0.001$; 2.7-kHz alternative: $P < 0.001$). The median latency required to meet

our choice criterion was 60 s (interquartile range 36–103 s; range 10–524 s).

In Experiment 2, phonotaxis scores differed significantly depending on the type of stimulus (Fig. 2b; $F_{5,155} = 35.5$, $P < 0.001$). Phonotaxis scores were significantly lower in response to both unimodal calls presented at both levels when compared with responses to the bimodal call (averaged over the first and last tests) (0.9 kHz, 82 dB SPL: $F_{1,31} = 42.3$, $P < 0.001$; 0.9 kHz, 85 dB SPL: $F_{1,31} = 99.3$, $P < 0.001$; 2.7 kHz, 82 dB SPL: $F_{1,31} = 445.2$, $P < 0.001$; 2.7 kHz, 85 dB SPL: $F_{1,31} = 359.2$, $P < 0.001$). Most subjects were unresponsive to unimodal calls. Five of 32 (15.6%) and 12 of 32 (37.5%) subjects responded to the 0.9-kHz stimulus presented at 82 and 85 dB SPL, respectively, and 4 of 32 (12.5%) and 3 of 32 (9.4%) subjects responded to the 2.7-kHz stimulus presented at 82 and 85 dB SPL, respectively. Averaged across both stimulus levels, phonotaxis

scores were significantly higher in response to the 0.9-kHz unimodal call compared with the 2.7-kHz unimodal call ($F_{1,31} = 5.8$, $P = 0.023$). Phonotaxis scores did not differ across the two tests of the bimodal stimulus ($F_{1,31} = 0.04$, $P = 0.849$).

Neural responses to unimodal and bimodal calls

Using the dichotomous decision tree depicted in Fig. 3, we classified units into one of 10 hierarchically organized categories based on differences in the mean numbers of spikes/stimulus elicited by the unimodal and bimodal calls expressed as a percentage of responses to the bimodal call. Based on visual inspections of the variability in responses across units, we adopted a tolerance level of ±10% for inclusion within a certain classification (see Fig. 3). This decision allowed us to classify a unit according to whether the magnitude of its response to one or both unimodal calls exceeded its response to the bimodal call by 10%, was between 90 and 110% of its response to the bimodal call, was between 10 and 90% of its response to the bimodal call, or was less than 10% of its response to the bimodal call.

Combination sensitive or combination insensitive?

The highest level of the hierarchy (Fig. 3a) classified units as either combination-sensitive (101/140, 72.1%) or combination-insensitive (39/140, 27.9%). Combination-sensitive units were those for which either the response to both unimodal calls was less than 90% of the response to the bimodal call or the response to one of the unimodal calls exceeded the response to the bimodal call by 10%. Hence, combination-sensitive units responded differently when the frequencies of 0.9 and 2.7 kHz occurred in combination in the bimodal call compared to when only one or the other of these two frequencies was presented alone in a unimodal call. A unit was classified as combination-insensitive when its response to one or both unimodal calls was within ±10% of its response to the bimodal call, and the maximum response to a unimodal call did not exceed the response to the bimodal call by 10%. Hence, combination-insensitive units were those units that responded similarly to combinations of 0.9 and 2.7 kHz compared with responses to one or the other frequency presented by itself.

Combination-sensitive units: enhancement or suppression?

At the second level of the hierarchy (Fig. 3b), combination-sensitive units could be divided into two subclasses based on whether responses to the bimodal call were enhanced or suppressed relative to responses to unimodal calls. Enhancement occurred when a combination-sensitive neuron's responses to the bimodal call exceed its responses to both unimodal calls (Fig. 4a–c), whereas suppression occurred when the neuron exhibited a lower response to the bimodal call compared to one of the unimodal calls (Fig. 4d–e). Each of these two subclasses of combination-sensitive units could be further subdivided at the third and lowest level of the hierarchy (Fig. 3c) depending on whether enhancement was due to facilitation, linear summation, or compressive summation, and whether suppression was elicited by 0.9 or 2.7 kHz. Figure 5a depicts the distribution of values computed for the index of interaction for all combination-sensitive units. This distribution was normally distributed (Kolmogorov–Smirnov test: $D = 0.088$, $P = 0.054$) around a mean (±SD, here and throughout) of -0.21 ± 0.24. The index of interaction differed significantly across subclasses of combination-sensitive units at the lowest level of the classification hierarchy ($F_{4,96} = 33.2$, $P < 0.001$; Fig. 5b).

Enhancement of responses to the bimodal call was observed in 28.6% (40/140) of units. Of these units, 22.5% (9/40), representing 6.4% (9/140) of all units and 8.9% (9/101) of combination-sensitive units, exhibited facilitation, which we defined as a response to the bimodal call that exceeded the linear sum of responses to the two unimodal calls by more than 10% (Fig. 4a). The median sum of these units' responses to the two unimodal calls was 75.4% of their response to the bimodal call. Units displaying facilitation had a mean index of interaction of $I = 0.25 \pm 0.30$, which was significantly greater than 0.0 ($t = 2.6$, df = 8, $P = 0.034$; Fig. 5b). Approximately half (4/9) of units exhibiting facilitation did not respond to one of the two unimodal calls (see individual traces in Fig. 4a), and the median of their response to the other unimodal call was 53.4% of their response to the bimodal call.

An additional 22.5% (9/40) of units displaying enhanced responses to the bimodal call, representing 8.9% (9/101) of combination-sensitive units and 6.4% (9/140) of all units, exhibited linear summation (Fig. 4b). For these units, the linear sum (median: 101.9%) of the magnitudes of their responses to the 0.9-kHz unimodal call (median: 66.2%) and the 2.7-kHz unimodal call (median: 30.8%) was within 10% of their response to the bimodal call. The mean index of interaction for units exhibiting linear summation was $I = -0.03 \pm 0.04$ and did not differ significantly from 0.0 ($t = -1.7$, df = 8, $P = 0.133$; Fig. 5b).

The remaining 55% (22/40) of units displaying enhanced responses to the bimodal call exhibited compressive summation. These units represented 15.7% (22/140) of all units and 21.8% (22/101) of combination-sensitive units. By definition, the linear sum of these units' responses to the two unimodal calls exceeded their response to the bimodal call by more than 10%. As shown in Fig. 4c, the median responses of these units to the 0.9-kHz and 2.7-kHz unimodal calls
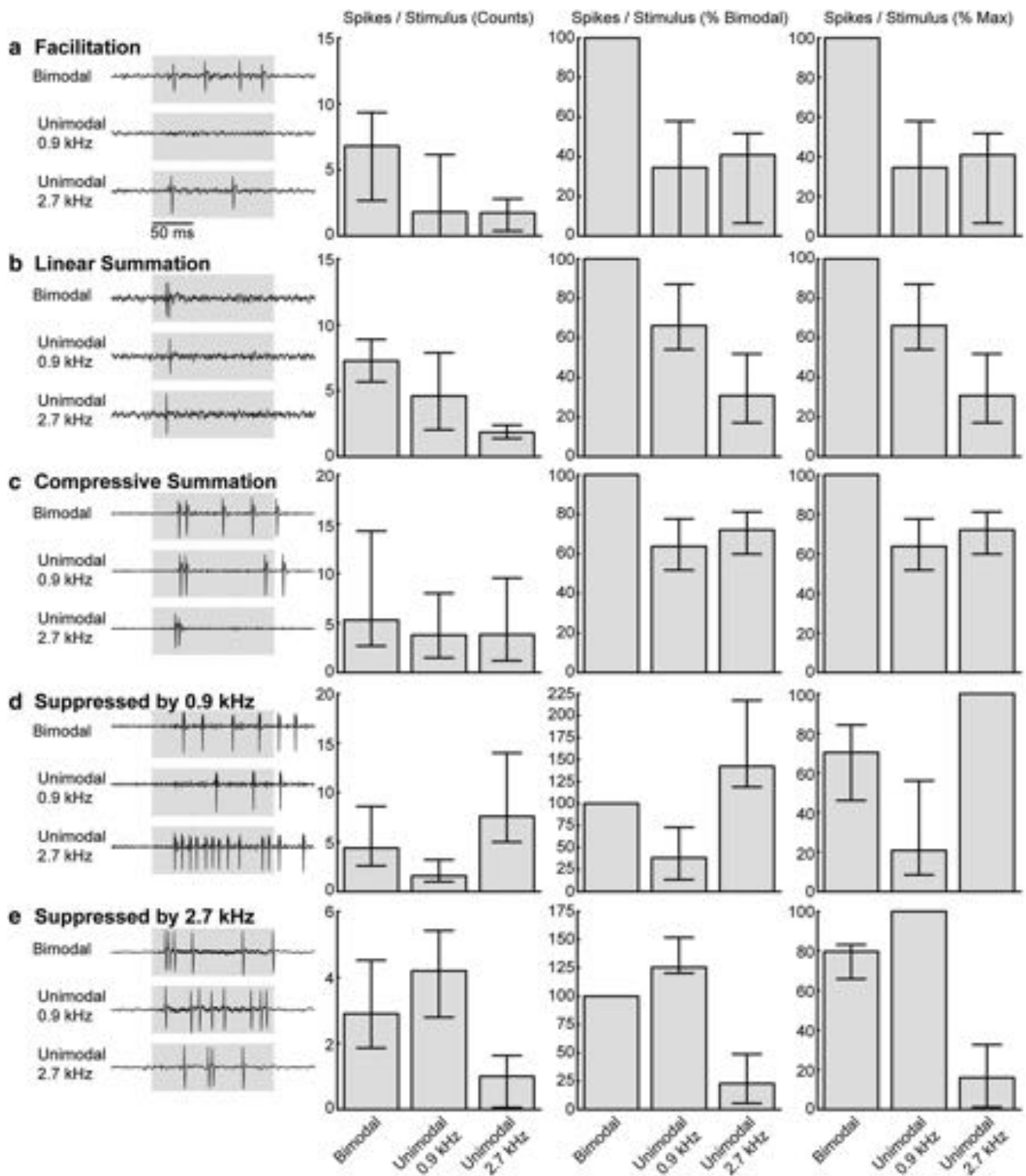
Fig. 4 Responses of combination-sensitive neurons to the bimodal call, the 0.9-kHz unimodal call, and the 2.7-kHz unimodal call. The *first column* shows representative neural traces of a selected unit of each type in response to one presentation of a bimodal call and each unimodal call. The shaded area corresponds to the timing of the 150-ms stimulus. The *second column* depicts the median (±interquartile range, IQR) number of spikes/stimulus. The *third column* displays the median (±IQR) number of spikes/stimulus expressed as a percentage of the response to the bimodal stimulus. The *fourth column* shows the median (±IQR) number of spikes/stimulus expressed as a percentage of the maximum stimulus-driven response. Data are shown separately for units classified according to the hierarchical classification scheme depicted in Fig. 3 as displaying **a** facilitation, **b** linear summation, **c** compressive summation, **d** suppression by the 0.9-kHz component, and **e** suppression by the 2.7-kHz component
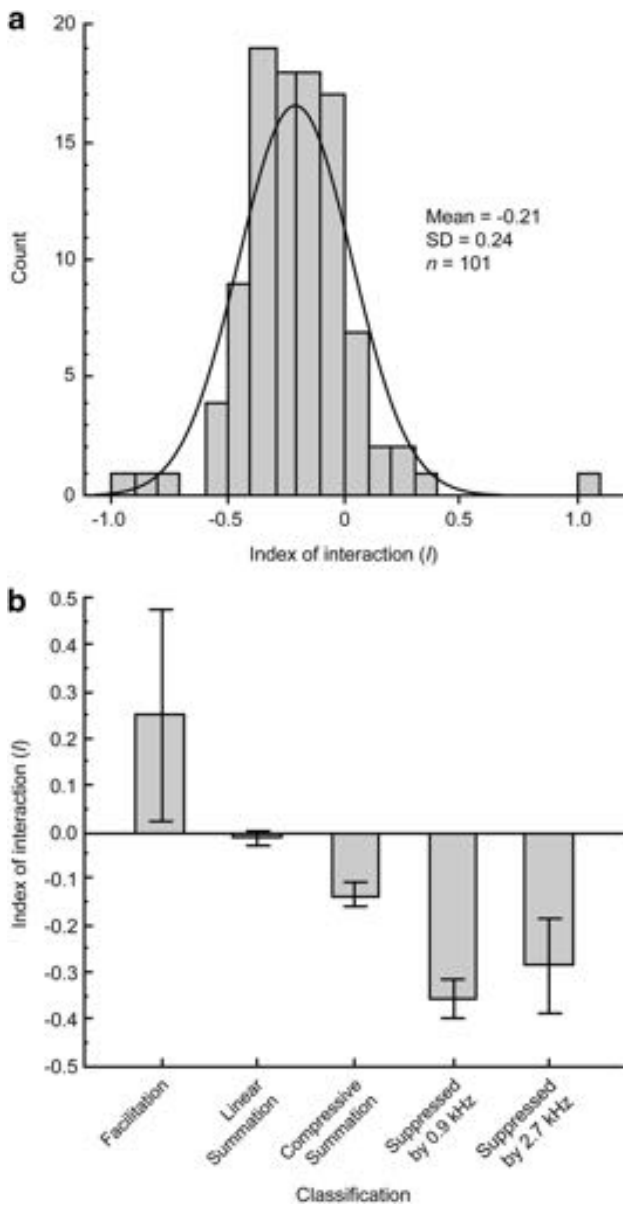
Fig. 5 Index of interaction. **a** Histogram showing the distribution of values computed for the index of interaction (*I*) across units. **b** Mean (±95% confidence intervals) values of the index of interaction for different subclasses of combination-sensitive units determined using the hierarchical classification scheme depicted in Fig. 3

were 63.8 and 72.1% of their responses to the bimodal call, respectively, yielding linearly summed responses to the two unimodal calls that exceeded their actual responses to the bimodal call by a median of 28.3%. Units exhibiting compressive summation had a mean index of interaction of $I = -0.14 \pm 0.06$, which was significantly less than 0.0 ($t = -10.2$, df $= 21$, $P < 0.001$; Fig. 5b).

In contrast to combination-sensitive units that exhibited enhanced responses to the bimodal call, 43.6% of all units (61/140), representing 60.4% of combination-sensitive

units, exhibited suppression of responses to the bimodal call compared with their responses to one or both unimodal calls (Fig. 4d, e). In these units, responses to one or both unimodal calls exceeded the response to the bimodal call by 10% or more. We found that 30.0% of all units (42/140) had maximal responses to the 2.7-kHz unimodal call and were suppressed by adding the 0.9-kHz component to the 2.7-kHz component in the bimodal call (Fig. 4d). The median response of these units to the 0.9-kHz and 2.7-kHz unimodal calls were 38.2 and 142.3% of their response to the bimodal call, respectively. For 8 of the 42 units suppressed by adding the 0.9-kHz component, the mean number of spikes per 150-ms stimulus recorded in response to the 0.9-kHz unimodal call fell within the 95% confidence interval around the mean spontaneous spike rate (i.e., ≤0.46 spikes/150 ms; see below). Responses of the remaining 34 of 42 units to the 0.9-kHz unimodal call exceeded the upper bound of the 95% confidence interval around the mean spontaneous rate. Units that were suppressed by the 0.9-kHz component had a mean index of interaction of $I = -0.35 \pm 0.15$, which was significantly less than 0.0 ($t = -15.4$, df $= 41$, $P < 0.001$; Fig. 5b).

We found that 13.6% of all units (19/140) were most responsive to the 0.9-kHz unimodal call and were suppressed by adding the 2.7-kHz component to the 0.9-kHz component in the bimodal call (Fig. 4e). The median responses of these units to the 0.9-kHz and 2.7-kHz unimodal calls were, respectively, 125.0 and 22.8% of their responses to the bimodal call. Nine of the 19 units suppressed by adding the 2.7-kHz component had responses to the 2.7-kHz unimodal call that fell within the 95% confidence interval around the mean spontaneous spike rate; responses for the remaining 10 of 19 units to the 2.7-kHz unimodal call exceeded the upper bound of the 95% confidence interval around the mean spontaneous rate. The mean index of interaction computed for units suppressed by adding the 2.7-kHz component was $I = -0.29 \pm 0.21$ and was significantly less than 0.0 ($t = -6.1$, df $= 18$, $P < 0.001$; Fig. 5b).

Combination-insensitive units: single-tone responders or multi-tone responders?

The 39 of 140 units (27.9%) classified as combination-insensitive units could be further subdivided into single-tone responders (STRs) and multi-tone responders (MTRs) at the second level of the classification hierarchy (Fig. 3b). STRs, which represented 6.4% of all units (9/140) and 23.1% of combination-insensitive units (9/39), were those units for which their response to one of the unimodal calls was within 10% of their response to the bimodal call, and their response to the other unimodal call was less than 10% of
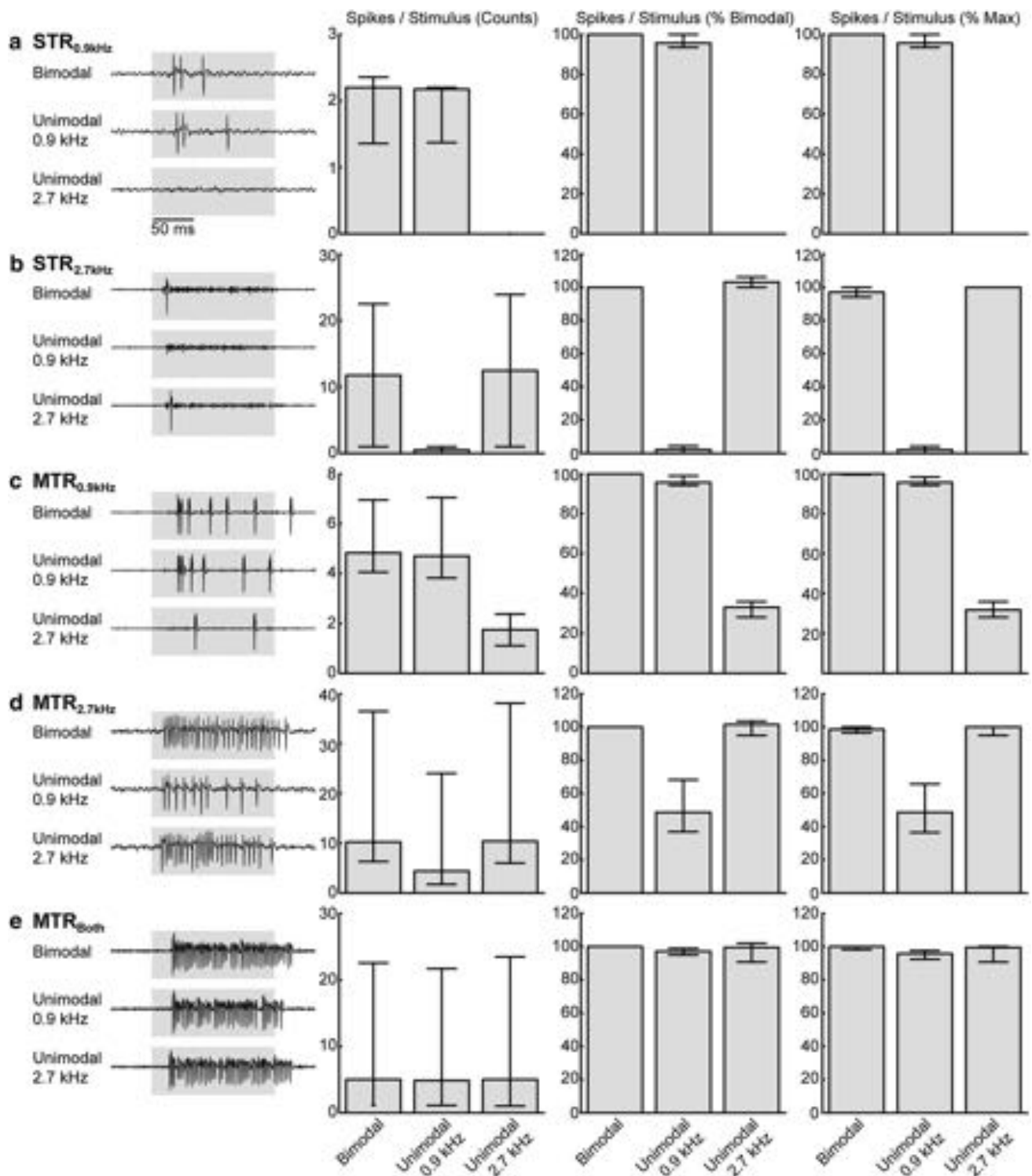
Fig. 6 Responses of combination-insensitive neurons to the bimodal call, the 0.9-kHz unimodal call, and the 2.7-kHz unimodal call. Details of individual plots are as described in the legend of Fig. 4. Data are shown separately for units classified according to the hierarchical classification scheme depicted in Fig. 3 as being **a** $STR_{0.9kHz}$ units, **b** $STR_{2.7kHz}$ units, **c** $MTR_{0.9kHz}$ units, **d** $MTR_{2.7kHz}$ units, and **e** MTR_{Both} units

Table 1 Summary of spiking activity in response to bimodal and unimodal calls

Response measure	Stimulus	Mean	Range	SD	n	Overall ANOVA			Contrast versus bimodal response		
						F	df	P	F	df	P
Mean spikes/stimulus	Bimodal	8.6	0.0–45.4	9.9	140	**26.8**	**2, 278**	**<0.001**	–	–	–
	0.9-kHz	5.4	0.0–36.3	7.4	140				**59.4**	**1, 139**	**<0.001**
	2.7-kHz	8.2	0.0–47.0	10.9	140				1.5	1, 139	0.222
Mean first-spike latency (ms)	Bimodal	24.8	2.1–115.6	19.7	139	**14.8**	**2, 234**	**<0.001**	–	–	–
	0.9-kHz	36.9	5.4–143.1	30.1	130				**43.0**	**1, 117**	**<0.001**
	2.7-kHz	28.2	2.6–129.6	26.0	127				**4.7**	**1, 117**	**0.031**
Mean inter-spike interval (ms)	Bimodal	20.7	3.3–114.5	16.4	134	**13.9**	**2, 182**	**<0.001**	–	–	–
	0.9-kHz	26.0	3.7–118.6	20.5	115				**33.9**	**1, 91**	**<0.001**
	2.7-kHz	16.7	2.6–85.2	16.7	111				1.0	1, 91	0.313
Variance in inter-spike interval	Bimodal	278.5	0.0–6110.0	617.3	134	1.7	2, 182	0.189	–	–	–
	0.9-kHz	325.6	0.0–2232.4	407.4	115				**12.3**	**1, 91**	**0.001**
	2.7-kHz	297.4	0.0–8678.5	871.9	111				2.0	1, 91	0.164

Significant differences are highlighted in bold text

their response to the bimodal call (Fig. 6a, b). In contrast, MTRs represented 21.4% of all units (30/140) and were classified as such because they responded well to frequencies of either 0.9 or 2.7 kHz in unimodal calls, but did not respond more when these two frequencies were combined in the bimodal call (Fig. 6c–e). MTRs exhibited a response to one of the unimodal calls that was within 10% of their response to the bimodal call, and their response to the other unimodal call was greater than 10% of their response to the bimodal call.[2]

Both STRs and MTRs were further subdivided at the third level of the hierarchy (Fig. 3c) based on their relative responses to the 0.9-kHz unimodal call ($STR_{0.9kHz}$ or $MTR_{0.9kHz}$) and 2.7-kHz unimodal call ($STR_{2.7kHz}$ or $MTR_{2.7kHz}$). Most (7/9, 77.8%) STR units, representing 5.0% of all units (7/140), were $STR_{0.9kHz}$ units that responded best to the 0.9-kHz unimodal call (Fig. 6a). The median responses of $STR_{0.9kHz}$ units to the 0.9-kHz and 2.7-kHz unimodal calls were 95.7 and 0.0% of their response to the bimodal call, respectively. $STR_{2.7kHz}$ units responded best to the 2.7-kHz unimodal call and represented just 1.4% (2/140) of all units (Fig. 6b). Their responses to the 0.9-kHz and 2.7-kHz unimodal calls were, respectively, 2.2 and 103.1% of their response to the bimodal call. Most MTR units could be further classified based on whether they responded best to the 0.9-kHz ($MTR_{0.9kHz}$) or 2.7-kHz ($MTR_{2.7kHz}$) unimodal call. $MTR_{0.9kHz}$ units, which represented 5.7% of all

units (8/140) and 26.7% of MTRs (8/30), had median responses to the 0.9-kHz and 2.7-kHz unimodal calls that were 95.7 and 33.0% of their responses to the bimodal call (Fig. 6c). In contrast, $MTR_{2.7kHz}$ units had median responses to the 0.9-kHz and 2.7-kHz unimodal calls that were 48.6 and 101.5% of their responses to the bimodal call (Fig. 6d). $MTR_{2.7kHz}$ units represented 12.1% of all units (17/140) and 56.7% of MTR units (17/30). For five additional MTR units (Fig. 6e), their responses to both unimodal calls were within 10% of their response to the bimodal call (0.9-kHz unimodal call: 97.1%; 2.7-kHz unimodal call: 99.5%). Hence, these MTR_{Both} units responded similarly to frequencies of 0.9 and 2.7 kHz presented separately, and responses were also similar when these two frequencies occurred in combination. They represented 16.7% of MTR units (5/30) and 3.6% of all units (5/140).

Classification based on spiking activity and frequency tuning

Most IC units exhibited little or no spontaneous activity. Across all 140 units, the median and modal spontaneous rates were both 0 spikes/s, with 74 of 140 units (52.9%) exhibiting no spontaneous activity, and 104 of 140 units (74.3%) exhibiting spontaneous firing rates less than 1 spike/s. Across all units, the mean spontaneous rate was 2.1 ± 5.7 spikes/s (95% confidence interval: 1.16 to 3.08 spikes/s, 0.17 to 0.46 spikes/150-ms epoch). For the 66 of 140 units (47.1%) with measurable spontaneous activity, the mean spontaneous rate was 3.8 ± 7.7 spikes/s. There were significant differences in spikes/stimulus, first-spike latency, and mean inter-spike interval in responses to the

---

[2] Note that our definition of MTR units differs from the use of this designation by Fuzessery and Feng (1982, 1983), who used it to refer to units that exhibited facilitated responses to combinations of multiple tones.
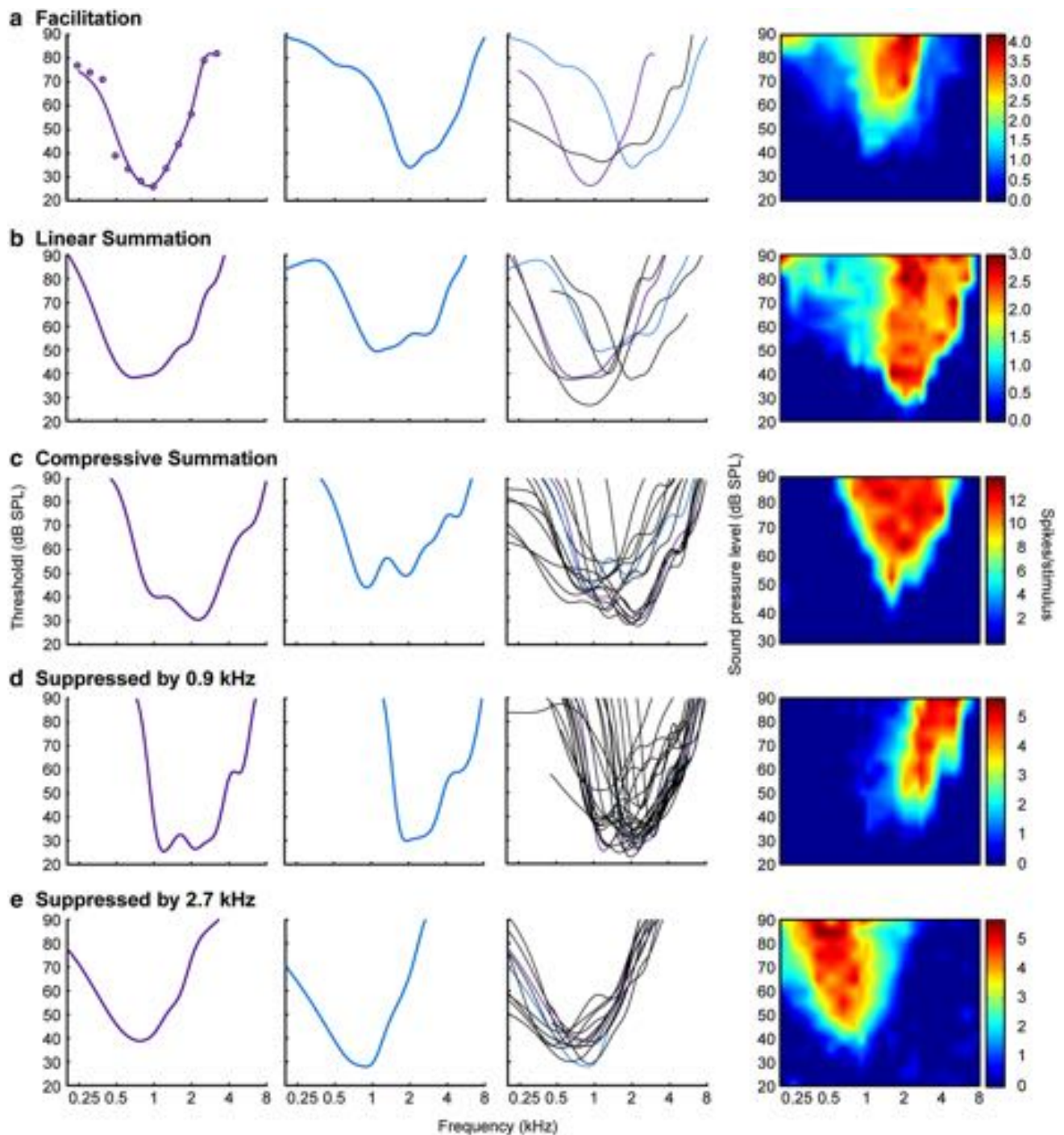
Fig. 7 Frequency tuning curves (FTCs) and frequency response maps for combination-sensitive units. The data shown here are from units classified according to the hierarchy depicted in Fig. 3 as **a** facilitation units, **b** linear summation units, **c** compressive summation units, **d** units suppressed by the 0.9-kHz component, and **e** units suppressed by the 2.7-kHz component. The first two columns depict FTCs from two representative units of the specified type. The plot in the first row and column shows the raw threshold values used to generate the fitted FTC in order to illustrate the average quality of a cubic spline fit. The individual R^2 value for this unit is 0.98, which is also the average over all 87 units for which FTCs were generated. The third column shows all FTCs determined for the specified type of unit. The fourth column shows a representative excitatory frequency response map from one unit of the specified type. Each representative FTC in the first and second column is depicted in the same respective *color* in the plot of all FTCs in the third column. The *color bar* depicts the number of spikes per stimulus for the frequency response map
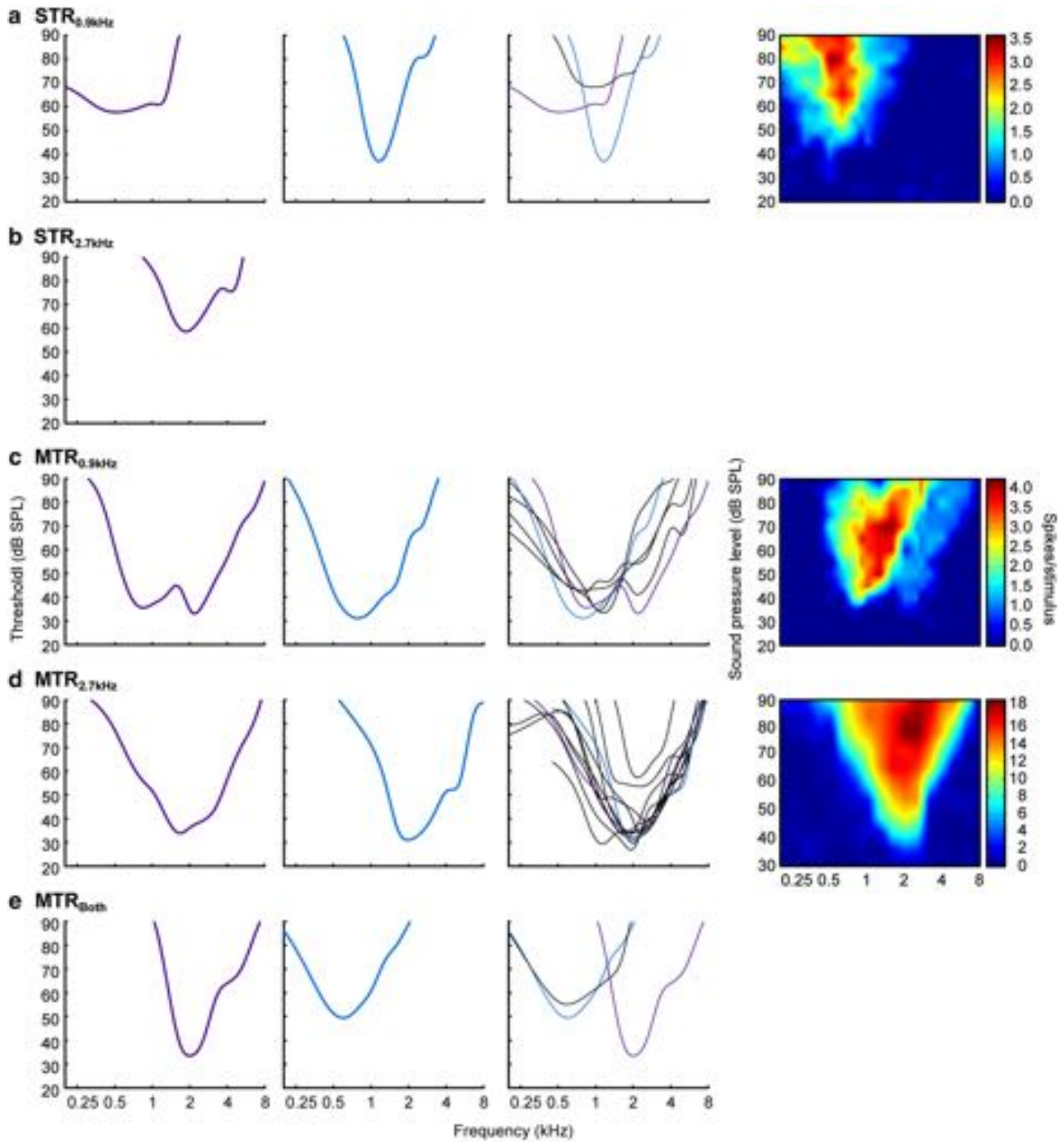
Fig. 8 Frequency tuning curves (FTCs) and frequency response maps for combination-insensitive units. The data shown here are from units classified according to the hierarchy depicted in Fig. 3 as

a $STR_{0.9kHz}$ units, **b** $STR_{2.7kHz}$ units, **c** $MTR_{0.9kHz}$ units, **d** $MTR_{2.7kHz}$ units, and **e** MTR_{Both} units. Details of columns are as described in the legend of Fig. 7

bimodal and unimodal calls (Table 1). Overall, the bimodal stimulus evoked more spikes/stimulus, shorter first-spike latencies, and intermediate inter-spike intervals compared with one or both of the unimodal stimuli (Table 1). The

variance in inter-spike interval was higher in response to the 0.9-kHz unimodal call compared with the bimodal call, but overall differences in this response measure were not significant (Table 1).

Table 2 Summary of properties measured from excitatory frequency tuning curves (FTCs)

Response measure	Mean	Minimum	Maximum	SD	n
Best excitatory frequency (Hz)	1503.1	454.0	3037.0	636.9	87
Threshold (dB SPL)	37.0	22.9	68.2	9.1	87
10-dB bandwidth (Hz)	1440.6	388.0	2751.9	613.4	86[b]
Q_{10dB}	1.2	0.5	3.8	0.6	86[b]
Low-mode best excitatory frequency (Hz)[a]	1129.4	454.0	1636.0	314.1	15
Low-mode threshold (dB SPL)[a]	36.9	25.4	55.2	7.0	15
High-mode best excitatory frequency (Hz)[a]	2168.1	1362.0	2521.0	292.5	15
High-mode threshold (dB SPL)[a]	39.0	26.4	59.4	10.5	15

[a] Measured for bimodal (W-shaped) FTCs

[b] Bandwidth could not be determined from one FTC because insufficiently low frequencies were tested

Figures 7 and 8 depict representative examples of FTCs and excitatory response maps for combination-sensitive and combination-insensitive units, respectively. Table 2 summarizes the properties we measured from the FTCs of 87 units. There were few clear patterns of association between the properties of FTCs and how units were classified. Averaging values of coefficients of variation (CV = 100% × SD/mean) across the four properties measured from FTCs (i.e., CV_{BEF}, $CV_{Threshold}$, $CV_{10-dB\ Bandwidth}$, and CV_{Q10dB}), units suppressed by the 2.7-kHz component (Fig. 7e) and $MTR_{2.7kHz}$ units (Fig. 8d) were the least variable (mean CVs of 21.2 and 22.1%, respectively), and $STR_{0.9kHz}$ units (Fig. 8a) and MTR_{Both} units (Fig. 8e) were the most variable (mean CV of 45.1 and 51.0%, respectively). Among combination-sensitive units, those exhibiting enhancement tended to be more variably tuned ($31.2\% \leq CV_{BEF} \leq 41.9\%$; Fig. 7a–c) than those displaying suppression ($20.0\% \leq CV_{BEF} \leq 26.9\%$; Fig. 7d, e). Across all 87 units, BEFs were distributed bimodally, with a low-frequency mode between 400 and 1500 Hz and a high-frequency mode between 1800 and 2600 Hz (Fig. 9a). The mean BEF, averaged over the 43 of 87 units (49.4%) with BEFs below 1500 Hz, was 924 ± 266 Hz; the mean BEF averaged over the 39 of 87 units (44.8%) with BEFs above 1800 Hz was 2121 ± 255 Hz. Units that responded best to one of the two unimodal calls tended to have BEFs in the corresponding mode. For example, units suppressed by 2.7 kHz (Fig. 7e) and $STR_{0.9kHz}$ units (Fig. 8a) both responded best to the 0.9-kHz component and had BEFs in the lower of the two modes. Relatively few units had BEFs near the grand mean of 1503 Hz reported in Table 2. Bimodality was also reflected in the tuning of a subset of 15 of 87 units (17.2%) with bimodal (W-shaped) FTCs (e.g., Figs. 7c, d, 8c). On average, units with bimodal FTCs had low-frequency and high-frequency modes centered around best frequencies of 1129 and 2168 Hz, respectively (Table 2). Across units, BEF was significantly negatively correlated with threshold (Fig. 9b, $r = -0.243$, $P = 0.023$, $n = 87$) and significantly positively correlated with 10-dB

bandwidth (Fig. 9c, $r = 0.504$, $P < 0.001$, $n = 86$) and Q_{10dB} values (Fig. 9d, $r = 0.472$, $P < 0.001$, $n = 86$).

An examination of frequency response maps revealed that some units were maximally driven by frequencies different from their BEF, and that this was related to the presence of inhibition. These units typically had regions of excitation that "tilted" toward frequencies higher or lower than their BEF as sound level was increased. Such units with BEFs above 1800 Hz were maximally driven by frequencies higher than their BEF (Figs. 7d, 10a), while those with BEFs below 1500 Hz were maximally driven by frequencies lower than their BEF (Figs. 7e, 10b). Figure 10 compares the excitatory and inhibitory frequency response maps for two such units classified as suppression units according to our classification hierarchy. Their maximal responses were shifted toward frequencies that were higher (Fig. 10a) or lower (Fig. 10b) than the BEF, and in both cases, excitatory regions were shifted away from a region of inhibition that was revealed using a two-tone inhibition paradigm. The examination of frequency response maps also revealed that some units (e.g., Fig. 8c) had "closed" tuning curves, meaning their response magnitudes at some frequencies first increased, and then decreased, as a function of increasing sound level, such that their maximal response occurred at intermediate sound levels.

We used DFAs to classify units as either combination-sensitive or combination-insensitive units (level 1), as enhanced, suppressed, STR, or MTR units (level 2), or into the 10 subclasses representing the lowest level of the classification hierarchy (level 3). None of the DFAs classified units as combination-sensitive or combination-insensitive units at levels significantly higher than expected by chance (Table 3). This is perhaps not surprising given that each of these two broad classes included units with quite different response properties (e.g., units displaying both enhancement and suppression are considered combination-sensitive units according to our hierarchy). The DFAs that used measures related to spike rate, spike timing, or all measures
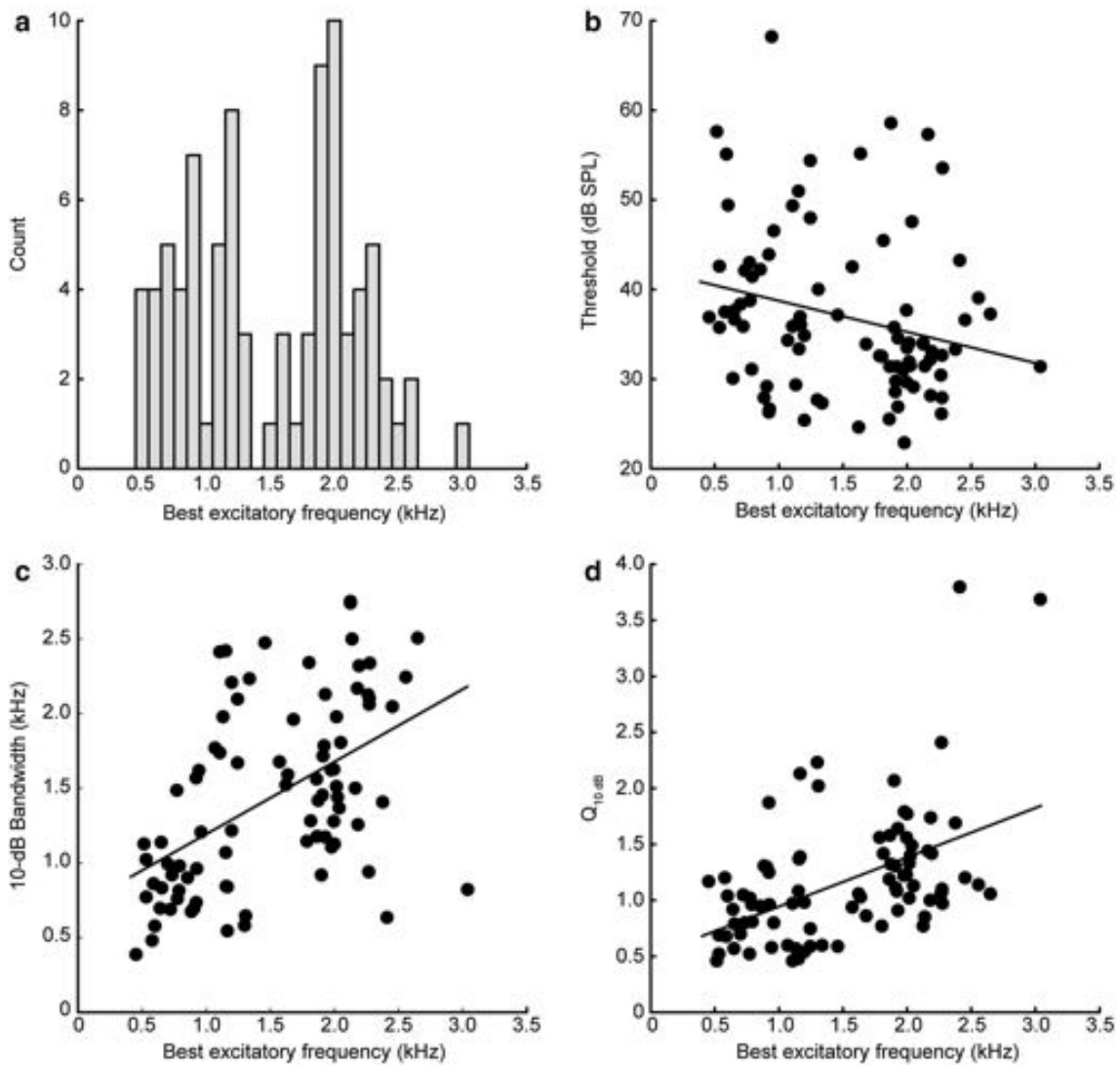
Fig. 9 Summary of values determined from frequency tuning curves (FTCs). **a** Histogram depicting the bimodal distribution of best excitatory frequencies (BEFs). **b** Relationship between threshold and BEF. **c** Relationship between 10-dB bandwidth and BEF. **d** Relationship between Q_{10dB} and BEF

together classified units as enhanced, suppressed, STR, or MTR units at levels significantly higher than expected by chance (Table 3). The DFA based on properties related to frequency tuning did not perform above chance levels at classifying units as enhanced, suppressed, STR, or MTR units. At the lowest level of the hierarchy, all four DFAs classified units into the 10 possible subclasses at above-chance levels (Table 3). The analysis using measures of spike rate as input performed best, with a classification success of 53.6% (corrected chance = 16.6%; kappa = 0.44; Table 3). Although several DFAs classified units at rates significantly higher than expected by chance, we would note that measures of classification success in these analyses were not particularly high, ranging between 37.9 and 59.3% (Table 3).

To further explore the potential importance of various response measures in correctly classifying units, we examined the standardized canonical discriminant function coefficients from the DFA that used all measures to classify units into the 10 subclasses at the lowest level of the hierarchy (Table 4). This analysis returned seven discriminant functions, the first three of which had eigenvalues greater than 1.0 and explained 85.5% of the variance in the input variables. The first discriminant function, which explained 43.3% of the variance, weighted the number of spikes/stimulus elicited by the bimodal and 2.7-kHz unimodal calls

more heavily than all other response measures. The number
of spikes/stimulus in responses to the 0.9-kHz and 2.7-kHz
unimodal calls were also heavily weighted by discrimination
function 3, which explained 15.2% of the variance. The sec-
ond discriminant function explained 27.0% of the variance
and weighted the mean and variance of inter-spike intervals
in responses to the 2.7-kHz unimodal call more heavily than
other variables. The mean inter-spike interval in response
to the 2.7-kHz unimodal call and BEF were also weighted
heavily by the third discriminant function.

Discussion

Two main results from our study can be summarized as
follows. First, in phonotaxis tests, females responded dif-
ferentially and selectively to bimodal calls compared with
unimodal calls. The combination of low-frequency and high-
frequency spectral components in a simulated advertisement
call was behaviorally more salient than either single com-
ponent presented alone. Second, a substantial proportion—
nearly three quarters—of the units we recorded in the IC
were combination sensitive in the sense that they displayed
either stronger or weaker responses to the spectral combi-
nation in the bimodal call compared with unimodal calls.
Of these combination-sensitive units, 40% exhibited either
facilitation (~9%), linear summation (~9%), or compressive
summation (~22%), and 60% exhibited suppression. With
the exception of units exhibiting linear summation, most
combination-sensitive units exhibited nonlinear responses.
Based on these results, we conclude that nonlinear processes
in the green treefrog's auditory midbrain play potentially
important roles in integrating information about the pre-
ferred multicomponent spectrum of the male's advertise-
ment call. Importantly, although the decision rules of our
hierarchical classification scheme assigned units to discrete
classes based on differences in their responses to bimodal
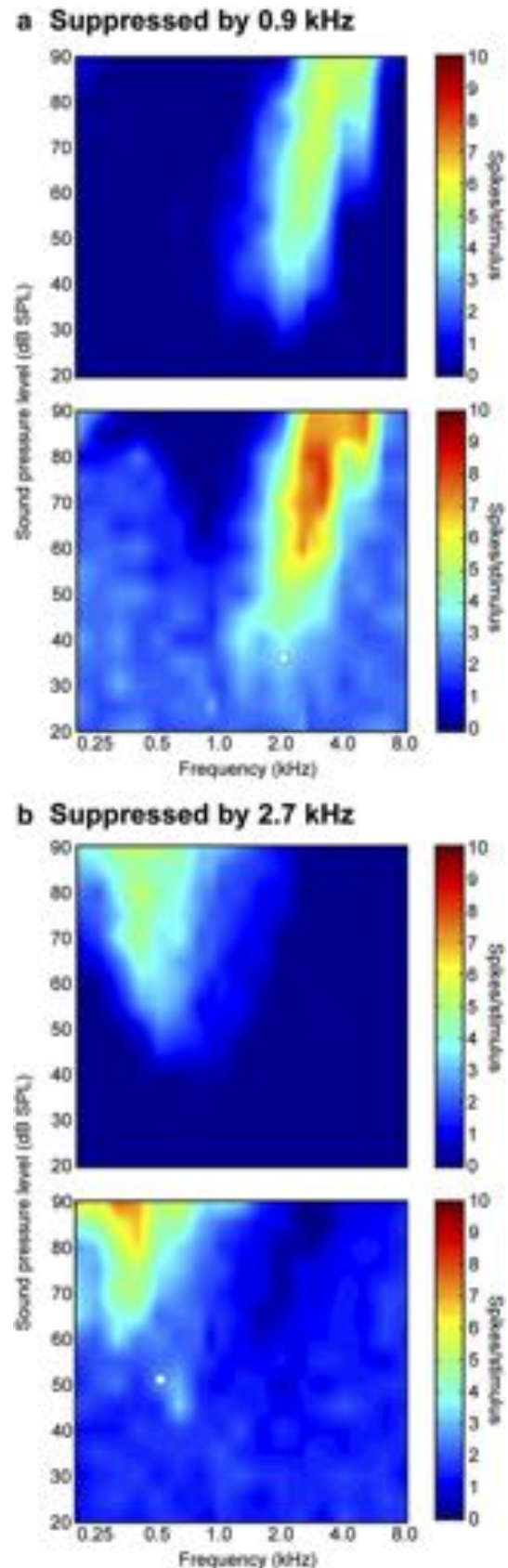and unimodal calls, there was actually continuous variation

a **Suppressed by 0.9 kHz**

b **Suppressed by 2.7 kHz**

Table 3 Comparison of classification successes from discriminant function analyses (DFA)

Hierarchy level	DFA input variables	Corrected chance (%)	Classification success (%)	Kappa	95% confidence interval	Z	P
Level 1[a]	Spike rate ($n = 140$)[d]	69.0	75.0	0.19	−0.04 to 0.43	1.54	0.062
	Spike timing ($n = 92$)[e]	71.8	69.6	−0.08	−0.42 to 0.26	−0.48	0.314
	Frequency tuning ($n = 86$)[f]	66.4	73.3	0.20	−0.08 to 0.49	1.33	0.090
	All ($n = 58$)[g]	62.5	70.7	0.22	−0.10 to 0.54	1.29	0.098
Level 2[b]	**Spike rate ($n = 140$)[d]**	**36.7**	**59.3**	**0.36**	**0.23 to 0.49**	**5.56**	**<0.001**
	Spike timing ($n = 92$)[e]	**33.9**	**54.3**	**0.31**	**0.15 to 0.47**	**4.14**	**<0.001**
	Frequency tuning ($n = 86$)[f]	33.8	33.7	0.00	−0.15 to 0.15	−0.02	0.490
	All ($n = 58$)[g]	**34.9**	**56.9**	**0.34**	**0.14 to 0.54**	**3.51**	**<0.001**
Level 3[c]	**Spike rate ($n = 140$)[d]**	**16.6**	**53.6**	**0.44**	**0.34 to 0.54**	**11.78**	**<0.001**
	Spike timing ($n = 92$)[e]	**16.6**	**38.0**	**0.26**	**0.14 to 0.38**	**5.53**	**<0.001**
	Frequency tuning ($n = 86$)[f]	**15.6**	**38.4**	**0.27**	**0.15 to 0.39**	**5.83**	**<0.001**
	All ($n = 58$)[g]	**13.0**	**37.9**	**0.29**	**0.14 to 0.43**	**5.64**	**<0.001**

Analyses with a classification success significantly higher than expected by chance are highlighted in bold text

[a] Combination-sensitive or combination-insensitive

[b] Enhancement, suppression, STR, or MTR

[c] Facilitation, linear summation, compressive summation, suppressed by 0.9 kHz, suppressed by 2.7 kHz, STR$_{0.9kHz}$, STR$_{2.7kHz}$, MTR$_{0.9kHz}$, MTR$_{2.7kHz}$, or MTR$_{Both}$

[d] "Spike rate" includes spontaneous rate (spikes/s) and responses to bimodal and unimodal calls (spikes/stimulus)

[e] "Spike timing" includes mean first-spike latency and mean and variance of inter-spike interval in responses to bimodal and unimodal calls

[f] "Frequency tuning" includes best excitatory frequency, threshold, 10-dB bandwidth, and Q$_{10dB}$

[g] "All" includes all variables included under spike rate, spike timing, and frequency tuning

Table 4 Standardized canonical discriminant function coefficients

Response measure	Stimulus	Function						
		1	2	3	4	5	6	7
Spontaneous rate (spikes/s)	None	−0.11	−0.72	0.64	−0.16	0.06	0.32	0.39
Mean spikes/stimulus	Bimodal	**−1.70**	0.94	0.48	**2.69**	**1.68**	0.67	−0.72
	0.9-kHz unimodal	−0.49	−0.53	**1.43**	−0.06	−0.85	0.33	0.06
	2.7-kHz unimodal	**2.69**	0.45	**−1.66**	**−2.59**	−0.71	−0.52	0.40
Mean first-spike latency (ms)	Bimodal	0.41	−0.70	−0.03	**1.56**	0.79	−0.04	0.23
	0.9-kHz unimodal	0.07	0.44	0.23	−0.69	−0.20	0.29	0.91
	2.7-kHz unimodal	−0.61	0.18	0.22	−0.53	**−1.29**	0.32	−0.24
Mean inter-spike interval (ms)	Bimodal	0.22	0.18	0.10	−0.59	−0.63	**−1.19**	−0.63
	0.9-kHz unimodal	0.40	−0.86	0.67	**1.66**	0.70	0.81	−0.56
	2.7-kHz unimodal	−0.31	**1.45**	**−1.26**	−0.83	0.76	0.03	0.86
Variance in inter-spike interval	Bimodal	0.42	−0.29	−0.15	0.58	0.29	0.82	0.80
	0.9-kHz unimodal	−0.20	0.57	−0.40	−0.57	−0.35	0.01	−0.19
	2.7-kHz unimodal	0.30	**−1.24**	0.61	0.22	0.34	−0.13	−0.36
Best excitatory frequency (Hz)	Pure tones	−0.15	−0.10	**1.09**	0.26	0.28	0.32	0.52
Threshold (dB SPL)	Pure tones	0.07	0.18	0.48	0.33	0.20	0.15	0.35
10-dB bandwidth (Hz)	Pure tones	0.27	0.60	−0.21	−0.01	−0.21	−0.76	−0.25
Q$_{10dB}$	Pure tones	0.24	0.51	−0.82	0.05	−0.21	−0.09	−0.28
Eigenvalue		3.3	2.1	1.2	0.6	0.3	0.2	0.1
Percentage of variance explained		43.3	27.0	15.2	7.4	3.6	2.7	0.8
Cumulative percentage of variance explained		43.3	70.3	85.5	92.9	96.5	99.2	100.0

Coefficients with absolute values greater than 1.0 are highlighted in bold text

between facilitation and suppression across the population of combination-sensitive units we recorded in the frog IC. These results reflect the known diversity of responses in the anuran IC. Although the DFAs classified units to the 10 subclasses of our classification hierarchy significantly above chance levels, the diversity of responses observed across subclasses of both combination-sensitive and combination-insensitive units accounts for the relatively low measures of classification success overall.

Multicomponent signals, peripheral tuning, and spectral preferences

It is not uncommon for anuran advertisement calls to have bimodal spectra in which simultaneous sound energy (e.g., harmonics or formant-like spectral components) occurs in two distinct regions of the frequency spectrum (Gerhardt and Huber 2002). Across anuran species, the two frequency regions emphasized in bimodal call spectra often closely match the species-specific tuning of the amphibian and basilar papillae in the inner ear, though this match is seldom perfect (Gerhardt and Schwartz 2001; Simmons 2013). This is also the case for green treefrogs (Moffat and Capranica 1974; Ehret and Capranica 1980). As in other frogs (Fuzessery 1988), auditory nerve fibers in the green treefrog can be divided into three classes (Moffat and Capranica 1974; Ehret and Capranica 1980). The low-frequency spectral component of advertisement calls is transduced by populations of hair cells of the amphibian papilla that are sensitive to low-frequency (e.g., <700 Hz) and mid-frequency (e.g., 700–1400 Hz) sounds. The high-frequency spectral component is transduced by the hair cells of the basilar papilla, which is sensitive to high-frequency (e.g., 2000–4000 Hz) sounds. In addition, studies using a reflex modification technique have shown that both the frequency sensitivity (Megela-Simmons et al. 1985) and frequency selectivity (Moss and Simmons 1986) of the green treefrog's hearing are best at the two frequency modes in the species' advertisement call. Hence, although the low-frequency and high-frequency spectral components of the advertisement call occur in a single sensory modality (audition), they nevertheless represent separate signal components that are processed through separate peripheral sensory channels. As noted earlier, the two components appear to serve somewhat different primary functions (Gerhardt 1976, 1981b; Rheinlaender et al. 1979; Klump et al. 2004).

In species that produce calls with bimodal spectra, such signals typically evoke more robust vocal responses from males (Capranica 1965, 1966) and phonotaxis from females (Gerhardt 1981a, 2005; Gerhardt et al. 2007; Bee 2010) compared with otherwise similar signals manipulated to have unimodal spectra. A series of earlier behavioral studies

of green treefrogs by Gerhardt (1974, 1976, 1981b, 1986) examined female preferences based on spectral content as a function of signal level, and many of our results corroborate his. For example, females tested in those studies generally preferred bimodal calls over both a unimodal call (i.e., a 0.9-kHz call) and bimodal calls in which one of the two spectral modes was attenuated, in some tests by as little as 6 dB. In the present study, females also exhibited robust behavioral selectivity for bimodal over unimodal calls. However, not all results are consistent between studies. Gerhardt (1974), for instance, reported that 70% of females responded when tested with a low-frequency unimodal call, and 5 of 5 responded to a high-frequency unimodal call. In contrast, females in the present study were generally unresponsive to the 0.9-kHz unimodal call, although they were significantly, albeit only slightly, more responsive to the 0.9-kHz unimodal call compared with the 2.7-kHz unimodal call. At present, we can only speculate about possible reasons underlying differences between our study and those of Gerhardt. In addition to potential differences between our study population (Texas) and his (Georgia), many of Gerhardt's tests were performed outdoors, over cement or low-cut grass, with a speaker separation 4 times greater than ours, and at sound levels well below those used in the present study (e.g., 45–75 dB SPL at the female's release site). Differences in the overall sound levels used in our study and those of Gerhardt seem particularly important given the general level-dependence of many of Gerhardt's findings. Our choice of levels (i.e., 85 dB SPL at the female release site for the bimodal call) was designed to simulate a male calling at close range near a female making a mate choice decision. Future integrative studies aimed at examining the level-dependence of both receiver behavior and the responses of IC neurons would be informative.

Multicomponent signals, spectral integration, and central auditory processing

Current evidence indicates that the central processing of bimodal calls transduced by the two inner-ear sensory papillae involves increasing neural selectivity for bimodal spectra at successive levels of the ascending auditory pathway (Hall 1994; Rose and Gooler 2007). Sensitivity to both of the frequency regions emphasized in advertisement calls appears to be absent in the dorsal medullary nucleus (the first processing stage in the anuran ascending auditory system), first emerges in the superior olivary nucleus, and is prominent by the level of the IC, where neurons exhibit considerable diversity in their frequency tuning. Our results suggest neurons in the green treefrog IC exhibit diversity in frequency tuning that is broadly similar to that reported for other anurans (Mudry et al. 1977; Walkowiak 1980; Fuzessery and Feng 1982; Hall 1999). In a single-unit study

of the IC in northern leopard frogs, for example, Fuzessery and Feng (1982) demonstrated that IC neurons display a diversity of sensitivity patterns for both frequency and intensity. Similar diversity of tuning was also observed in the green treefrog IC, and this diversity can account for the limited contribution of properties related to frequency tuning in classifying units in DFAs according to our hierarchical classification scheme. Fuzessery and Feng (1982) also reported that many IC neurons had tilted FTCs and iso-response contours, such that maximal firing rates occurred at frequencies different from the BEF. Such tuning appears to be tied closely to two-tone inhibition, which was a common feature of IC neurons in the study by Fuzessery and Feng (1982) that was critical in determining the shape of a neuron's excitatory receptive field. Our excitatory frequency response maps revealed similarly tilted regions of excitation in the tuning of IC neurons in green treefrogs that also appeared to be shaped by inhibition. Fuzessery and Feng (1982) also reported that two-tone inhibition by relatively lower frequencies was somewhat more common than inhibition by relatively higher frequencies. Consistent with their observations, we found that units suppressed by the 0.9-kHz component were more than twice as likely to be encountered as those suppressed by the 2.7-kHz component. Relatively little of the two-tone inhibition observed in the rate responses of anuran IC neurons appears to be accounted for by the two-tone rate suppression observed in the responses of low-frequency-tuned auditory nerve fibers suppressed by mid-frequencies due to known mechanical interactions along the amphibian papilla (Fuzessery and Feng 1982). Hall (1999) later confirmed using bicuculline to block $GABA_A$ receptors that GABAergic inhibition in the IC of northern leopard frogs shapes frequency selectivity. Similar inhibitory mechanisms likely account for the 28% (17/61) of suppression units reported in the present study having responses to one of the unimodal calls that did not exceed the upper-bound 95% confidence interval around the mean spontaneous rate. That is, these units did not exhibit a measurable excitatory response to one of the unimodal calls, but the presence of the corresponding spectral component in the bimodal call resulted in suppression compared to their responses to the other unimodal call. Most suppression units, however, exhibited excitatory responses to both unimodal calls that exceeded the upper-bound 95% confidence interval around the mean spontaneous rate but nevertheless displayed a suppressed response to the bimodal call compared with the most salient unimodal call. The precise role of inhibition in these neurons' suppressed responses to bimodal calls requires further investigation.

At the level of the auditory thalamus, stimuli broadly simulating bimodal calls produce facilitated responses that are greater than the sum of responses to the two corresponding unimodal stimuli (Fuzessery and Feng 1983; Megela 1983; Mudry and Capranica 1987a, b). Evidence that facilitation in response to bimodal spectra may occur prior to the auditory thalamus was first reported by Fuzessery and Feng (1982). In that study of northern leopard frogs, 21 of 130 IC units (16%) responded best to combinations of two or more simultaneous tones. Most of these facilitation units were driven best by two tones separated by relatively small frequency differences (e.g., <200 Hz). Just two neurons (2/130, 1.5%) were reported to respond only to widely separated frequencies that were transduced by the amphibian and basilar papillae, similar to the bimodal calls used in our study. These findings were inconsistent with results from an earlier study of evoked potentials in the IC of the same species, which had shown linear summation but not facilitation to combination tones (Mudry et al. 1977). Unfortunately, no data exist describing whether northern leopard frogs behaviorally discriminate between signals with bimodal versus unimodal spectra, making comparisons between behavior and neurophysiology much more indirect for that species.

In the present study, bimodal calls were behaviorally more salient, and 28.6% of recorded IC neurons displayed an enhanced response to the combination of frequencies in the same bimodal call, with 6.4% of units exhibiting facilitation, 6.4% exhibiting linear summation, and 15.7% exhibiting compressive summation. Our results, therefore, corroborate and extend earlier work on frogs by confirming that facilitated responses to combinations of widely separated frequencies are indeed found in low proportions in the IC of an anuran species that is behaviorally selective for calls with bimodal spectra. An important finding to emphasize from the present study, however, is that while the frequency combination in the bimodal call elicited greater responses in behavior and from some combination-sensitive units, this was not the case for a majority (71.4%) of IC neurons, for which we saw no evidence for enhancement of responses to the bimodal call. Instead, most IC neurons exhibited responses to bimodal calls that either were the same as their response to one of the unimodal calls (STRs and MTRs) or were suppressed relative to their responses to one of the unimodal calls. How various combination-sensitive and combination-insensitive neurons in the IC might interact locally or feed forward to, or receive feedback from, targets in the auditory thalamus and forebrain to effect preferences for bimodal calls is presently unknown and deserves further study. It will also be important to understand how a female's reproductive state influences the responses of single neurons in the green treefrog's IC (Goense and Feng 2005; Miranda and Wilczynski 2009).

Combination-sensitive neurons and the processing of multicomponent signals

In addition to anurans, combination-sensitive processing of spectral information in hearing and sound communication has been demonstrated in songbirds (Margoliash and Fortune 1992), primates (Kadia and Wang 2003), rodents (Akimov et al. 2017), and bats (Fitzpatrick et al. 1993). Most of the previous work on combination-sensitive processing of the acoustic spectra of communication sounds has been done in the context of bat echolocation, where neurons in the IC (Mittmann and Wenstrup 1995; Yan and Suga 1996; Nataraj and Wenstrup 2006), auditory thalamus (Olsen and Suga 1991a, b), and auditory cortex (Suga et al. 1983; Kawasaki et al. 1988; Fitzpatrick et al. 1993) encode critical biosonar information about target identity, velocity, and distance through combination-sensitive spectral and temporal analyses of pulse–echo pairs. While combination-sensitive units in the bat auditory system appear highly specialized for echolocation, similar processing undoubtedly underlies perception of diverse communication signals, including speech (Sussman et al. 1998). We suggest facilitation, summation, and suppression are general combination-sensitive mechanisms in sensory processing that are likely to play critical roles in the perceptual analysis of multicomponent and multimodal signals by receivers in a broad diversity of taxa. Our work highlights the potential for both linear processes (e.g., linear summation) and especially nonlinear processes (e.g., facilitation, compressive summation, and suppression) to contribute mechanistically toward a receiver's perception of multicomponent signals. These mechanisms are likely to be generally important in the context of perceptual grouping (Medvedev et al. 2002; Farris and Taylor 2016). For example, in the context of auditory scene analysis (Bregman 1990), combination-sensitive neurons that are also temporally selective for synchrony among the onsets or offsets of simultaneous components of the frequency spectrum likely contribute to the effectiveness of temporal coherence (Elhilali et al. 2009; Lu et al. 2017) as an auditory grouping cue. Integrating future research on the perception of multicomponent and multimodal communication signals, the acoustic cues that promote perceptual grouping, and combination-sensitive neuronal processing would significantly advance our understanding of the neural mechanisms underlying animal communication.

References

Akimov AG, Egorova MA, Ehret G (2017) Spectral summation and facilitation in on- and off-responses for optimized representation of communication calls in mouse inferior colliculus. Eur J Neurosci 46(3):440–459

Bee MA (2010) Spectral preferences and the role of spatial coherence in simultaneous integration in gray treefrogs (*Hyla chrysoscelis*). J Comp Psychol 124:412–424

Bee MA, Miller CT (2016) Psychological mechanisms in animal communication, vol 5. Animal signals and communication. Springer, Berlin

Bradbury JW, Vehrencamp SL (2011) Principles of animal communication, 2nd edn. Sinauer Associates, Sunderland

Bregman AS (1990) Auditory scene analysis: the perceptual organization of sound. M.I.T. Press, Cambridge

Bro-Jørgensen J (2010) Dynamics of multiple signalling systems: animal communication in a world in flux. Trends Ecol Evol 25(5):292–300

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

Candolin U (2003) The use of multiple cues in mate choice. Biol Rev 78(4):575–595

Capranica RR (1965) The evoked vocal response of the bullfrog: a study of communication by sound. M.I.T. Press, Cambridge

Capranica RR (1966) Vocal response of the bullfrog to natural and synthetic mating calls. J Acoust Soc Am 40:1131–1139

Ehret G, Capranica RR (1980) Masking patterns and filter characteristics of auditory nerve fibers in the green treefrog (*Hyla cinerea*). J Comp Physiol 141(1):1–12

Elhilali M, Ma L, Micheyl C, Oxenham AJ, Shamma SA (2009) Temporal coherence in the perceptual organization and cortical representation of auditory scenes. Neuron 61(2):317–329

Farris HE, Taylor RC (2016) Mate searching animals as model systems for understanding perceptual grouping. In: Bee MA, Miller CT (eds) Psychological mechanisms in animal communication, vol 5. Animal signals and communication. Springer, New York, pp 89–118

Fitzpatrick DC, Kanwal JS, Butman JA, Suga N (1993) Combination-sensitive neurons in the primary auditory cortex of the mustached bat. J Neurosci 13(3):931–940

Fuzessery ZM (1988) Frequency tuning in the anuran central auditory system. In: Fritzsch B, Ryan MJ, Wilczynski W, Hetherington TE, Walkowiak W (eds) The evolution of the amphibian auditory system. Wiley, New York, pp 253–273

Fuzessery ZM, Feng AS (1982) Frequency selectivity in the anuran auditory midbrain: single unit responses to single and multiple tone stimulation. J Comp Physiol 146(4):471–484

Fuzessery ZM, Feng AS (1983) Mating call selectivity in the thalamus and midbrain of the leopard frog (*Rana p. pipiens*): single and multiunit analyses. J Comp Physiol 150(3):333–344

Gerhardt HC (1974) The significance of some spectral features in mating call recognition in the green treefrog (*Hyla cinerea*). J Exp Biol 61(1):229–241

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

Gerhardt HC (1976) Significance of two frequency bands in long distance vocal communication in the green treefrog. Nature 261(5562):692–694

Gerhardt HC (1981a) Mating call recognition in the barking treefrog (*Hyla gratiosa*): responses to synthetic calls and comparisons with the green treefrog (*Hyla cinerea*). J Comp Physiol 144(1):17–25

Gerhardt HC (1981b) Mating call recognition in the green treefrog (*Hyla cinerea*): importance of two frequency bands as a function of sound pressure level. J Comp Physiol 144(1):9–16

Gerhardt HC (1986) Recognition of spectral patterns in the green treefrog: neurobiology and evolution. Exp Biol 45(3):167–178

Gerhardt HC (1992) Multiple messages in acoustic signals. Semin Neurosci 4:391–400

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 Verlag, Basel, pp 209–220

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

Gerhardt HC (2005) Acoustic spectral preferences in two cryptic species of grey treefrogs: implications for mate choice and sensory mechanisms. Anim Behav 70:39–48

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

Gerhardt HC, Schwartz JJ (2001) Auditory tuning, frequency preferences and mate choice in anurans. In: Ryan MJ (ed) Anuran communication. Smithsonian Institution Press, Washington DC, pp 73–85

Gerhardt HC, Martinez-Rivera CC, Schwartz JJ, Marshall VT, Murphy CG (2007) Preferences based on spectral differences in acoustic signals in four species of treefrogs (Anura: Hylidae). J Exp Biol 210(17):2990–2998

Girard MB, Elias DO, Kasumovic MM (2015) Female preference for multi-modal courtship: multiple signals are important for male mating success in peacock spiders. Proc R Soc B 282(1820):20152222

Goense JBM, Feng AS (2005) Seasonal changes in frequency tuning and temporal processing in single neurons in the frog auditory midbrain. J Neurobiol 65(1):22–36

Hall JC (1994) Central processing of communication sounds in the anuran auditory system. Am Zool 34(6):670–684

Hall JC (1999) GABAergic inhibition shapes frequency tuning and modifies response properties in the auditory midbrain of the leopard frog. J Comp Physiol A 185(5):479–491

Hebets EA, Papaj DR (2005) Complex signal function: developing a framework of testable hypotheses. Behav Ecol Sociobiol 57(3):197–214

Higham JP, Hebets EA (2013) An introduction to multimodal communication. Behav Ecol Sociobiol 67(9):1381–1388

Kadia SC, Wang XQ (2003) Spectral integration in AI of awake primates: neurons with single- and multipeaked tuning characteristics. J Neurophysiol 89(3):1603–1622

Kawasaki M, Margoliash D, Suga N (1988) Delay-tuned combination-sensitive neurons in the auditory cortex of the vocalizing mustached bat. J Neurophysiol 59(2):623–635

Klump GM, Benedix JH, 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(12):1011–1021

Lee N, Ward JL, Vélez A, Micheyl C, Bee MA (2017) Frogs exploit statistical regularities in noisy acoustic scenes to solve cocktail-party-like problems. Curr Biol 27(5):743–750

Leroy SA, Wenstrup JJ (2000) Spectral integration in the inferior colliculus of the mustached bat. J Neurosci 20(22):8533–8541

Lu K, Xu YB, Yin PB, Oxenham AJ, Fritz JB, Shamma SA (2017) Temporal coherence structure rapidly shapes neuronal interactions. Nat Commun 8:13900

Margoliash D, Fortune ES (1992) Temporal and harmonic combination-sensitive neurons in the zebra finch's HVc. J Neurosci 12(11):4309–4326

Medvedev AV, Chiao F, Kanwal JS (2002) Modeling complex tone perception: grouping harmonics with combination-sensitive neurons. Biol Cybern 86(6):497–505

Megela AL (1983) Auditory response properties of the anuran thalamus: nonlinear facilitation. In: Ewert JP, Capranica RR, Ingle DJ (eds) Advances in vertebrate neuroethology. Plenum Press, New York, pp 895–899

Megela-Simmons A, Moss CF, Daniel KM (1985) Behavioral audiograms of the bullfrog (*Rana catesbeiana*) and the green tree frog (*Hyla cinerea*). J Acoust Soc Am 78(4):1236–1244

Milinski M, Bakker TC (1990) Female sticklebacks use male coloration in mate choice and hence avoid parasitized males. Nature 344(6264):330

Miranda JA, Wilczynski W (2009) Female reproductive state influences the auditory midbrain response. J Comp Physiol A 195(4):341–349

Mittmann DH, Wenstrup JJ (1995) Combination-sensitive neurons in the inferior colliculus. Hear Res 90(1–2):185–191

Moffat AJM, Capranica RR (1974) Sensory processing in the peripheral auditory system of treefrogs (*Hyla*). J Acoust Soc Am 55(2):480

Moss CF, Simmons AM (1986) Frequency selectivity of hearing in the green treefrog, *Hyla cinerea*. J Comp Physiol A 159(2):257–266

Mudry KM, Capranica RR (1987a) Correlation between auditory evoked responses in the thalamus and species-specific call characteristics I. *Rana catesbeiana* (Anura, Ranidae). J Comp Physiol A 160(4):477–489

Mudry KM, Capranica RR (1987b) Correlation between auditory thalamic area evoked responses and species-specific call characteristics II. *Hyla cinerea* (Anura: Hylidae). J Comp Physiol A 161(3):407–416

Mudry KM, Constantine-Paton M, Capranica RR (1977) Auditory sensitivity of the diencephalon of the leopard frog (*Rana p. pipiens*). J Comp Physiol 114(1):1–13

Nataraj K, Wenstrup JJ (2006) Roles of inhibition in complex auditory responses in the inferior colliculus: inhibited combination-sensitive neurons. J Neurophysiol 95(4):2179–2192

Nieuwenhuys R, ten Donkelaar HJ, Nicholson C (1998) The central nervous system of vertebrates. Springer, Heidelberg

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

Olsen JF, Suga N (1991a) Combination-sensitive neurons in the medial geniculate body of the mustached bat: encoding of relative velocity information. J Neurophysiol 65(6):1254–1274

Olsen JF, Suga N (1991b) Combination-sensitive neurons in the medial geniculate body of the mustached bat: encoding of target range information. J Neurophysiol 65(6):1275–1296

Partan SR (2013) Ten unanswered questions in multimodal communication. Behav Ecol Sociobiol 67(9):1523–1539

Partan SR, Marler P (1999) Communication goes multimodal. Science 283(5406):1272–1273

Partan SR, Marler P (2005) Issues in the classification of multimodal communication signals. Am Nat 166(2):231–245

Rheinlaender J, Gerhardt HC, Yager DD, Capranica RR (1979) Accuracy of phonotaxis by the green treefrog (*Hyla cinerea*). J Comp Physiol 133(4):247–255

Rose GJ, Gooler DM (2007) Function of the amphibian central auditory system. In: Narins PA, Feng AS, Fay RR, Popper AN (eds) Hearing and sound communication in amphibians, vol 28. Springer handbook of auditory research. Springer, New York, pp 250–290

Rowe C (1999) Receiver psychology and the evolution of multicomponent signals. Anim Behav 58:921–931

Rowe C, Halpin C (2013) Why are warning displays multimodal? Behav Ecol Sociobiol 67(9):1425–1439

Rowland WJ (1994) Proximate determinants of stickleback behaviour: an evolutionary perspective. In: Bell MA, Foster SA (eds) The evolutionary biology of the threespine stickleback. Oxford University Press, Oxford, pp 297–344

Ryan MJ (1985) The túngara frog: a study in sexual selection and communication. Chicago University Press, Chicago

Simmons AM (2013) "To ear is human, to frogive is divine": Bob Capranica's legacy to auditory neuroethology. J Comp Physiol A 199(3):169–182

Suga N, Oneill WE, Kujirai K, Manabe T (1983) Specificity of combination-sensitive neurons for processing of complex biosonar signals in auditory cortex of the mustached bat. J Neurophysiol 49(6):1573–1626

Sussman HM, Fruchter D, Hilbert J, Sirosh J (1998) Linear correlates in the speech signal: the orderly output constraint. Behav Brain Sci 21(02):241–259

Titus K, Mosher JA, Williams BK (1984) Chance-corrected classification for use in discriminant-analysis: ecological applications. Am Midl Nat 111(1):1–7

Uetz GW, Roberts JA, Clark DL, Gibson JS, Gordon SD (2013) Multimodal signals increase active space of communication by wolf spiders in a complex litter environment. Behav Ecol Sociobiol 67(9):1471–1482

Walkowiak W (1980) The coding of auditory signals in the torus semicircularis of the fire-bellied toad and the grass frog: responses to simple stimuli and to conspecific calls. J Comp Physiol 138(2):131–148

Ward JL, Buerkle NP, Bee MA (2013a) Spatial release from masking improves sound pattern discrimination along a biologically relevant pulse-rate continuum in gray treefrogs. Hear Res 306:63–75

Ward JL, Love EK, Vélez A, Buerkle NP, O'Bryan LR, Bee MA (2013b) Multitasking males and multiplicative females: dynamic signalling and receiver preferences in Cope's grey treefrog. Anim Behav 86(2):231–243

Wootton RJ (1976) Biology of the sticklebacks. Academic Press, London

Yan J, Suga N (1996) The midbrain creates and the thalamus sharpens echo-delay tuning for the cortical representation of target-distance information in the mustached bat. Hear Res 93(1–2):102–110