

Lars Holmstrom, Patrick D. Roberts and Christine V. Portfors

J Neurophysiol 98:3461-3472, 2007. First published Oct 10, 2007; doi:10.1152/jn.00638.2007

You might find this additional information useful...

This article cites 33 articles, 16 of which you can access free at:

<http://jn.physiology.org/cgi/content/full/98/6/3461#BIBL>

Updated information and services including high-resolution figures, can be found at:

<http://jn.physiology.org/cgi/content/full/98/6/3461>

Additional material and information about *Journal of Neurophysiology* can be found at:

<http://www.the-aps.org/publications/jn>

This information is current as of February 3, 2008 .

Responses to Social Vocalizations in the Inferior Colliculus of the Mustached Bat Are Influenced by Secondary Tuning Curves

Lars Holmstrom,¹ Patrick D. Roberts,² and Christine V. Portfors³

¹Systems Science Program, Portland State University, Portland, Oregon; ²Neurological Sciences Institute, Oregon Health and Science University, Beaverton, Oregon; and ³School of Biological Sciences, Washington State University, Vancouver, Washington

Submitted 9 June 2007; accepted in final form 1 October 2007

Holmstrom L, Roberts PD, Portfors CV. Responses to social vocalizations in the inferior colliculus of the mustached bat are influenced by secondary tuning curves. *J Neurophysiol* 98: 3461–3472, 2007. First published October 10, 2007; doi:10.1152/jn.00638.2007. Neurons in the inferior colliculus (IC) of the mustached bat integrate input from multiple frequency bands in a complex fashion. These neurons are important for encoding the bat's echolocation and social vocalizations. The purpose of this study was to quantify the contribution of complex frequency interactions on the responses of IC neurons to social vocalizations. Neural responses to single tones, two-tone pairs, and social vocalizations were recorded in the IC of the mustached bat. Three types of data driven stimulus-response models were designed for each neuron from single tone and tone pair stimuli to predict the responses of individual neurons to social vocalizations. The first model was generated only using the neuron's primary frequency tuning curve, whereas the second model incorporated the entire hearing range of the animal. The extended model often predicted responses to many social vocalizations more accurately for multiply tuned neurons. One class of multiply tuned neuron that likely encodes echolocation information also responded to many of the social vocalizations, suggesting that some neurons in the mustached bat IC have dual functions. The third model included two-tone frequency tunings of the neurons. The responses to vocalizations were better predicted by the two-tone models when the neuron had inhibitory frequency tuning curves that were not near the neuron's primary tuning curve. Our results suggest that complex frequency interactions in the IC determine neural responses to social vocalizations and some neurons in IC have dual functions that encode both echolocation and social vocalization signals.

INTRODUCTION

Many animal species have complex social structures that remain stable through the use of social communication signals. In many species, these signals are acoustic. To correctly interpret an acoustic signal to output an appropriate motor response, the receiver must detect, discriminate, and categorize each particular social vocalization. Presumably these tasks are performed by auditory neurons in the central auditory system, but how social vocalizations are encoded at different levels of the ascending auditory system is not well understood.

Few studies have examined encoding of social vocalizations in the auditory midbrain, particularly in mammalian species. The inferior colliculus (IC) is a midbrain structure that is a major relay station for both ascending and descending auditory pathways (Adams 1979). The IC has been well studied in several mammalian species and contains a tonotopic map

where low frequencies are represented dorsolaterally and high frequencies represented ventromedially (Lippe and Rubel 1983; Zook et al. 1985). In addition, many neurons in the IC are tuned to additional frequencies that are not predicted based on the tonotopic location of the neuron; they are multiply tuned (Portfors and Wenstrup 2002; Portfors and Felix 2005). Other neurons show facilitated or inhibited responses (combination sensitivity) to the combination of distinct frequency elements in complex sounds (Mittmann and Wenstrup 1995; Portfors and Wenstrup 1999). These response features are indicative of spectral integration at the level of single neurons and suggest that there may be neural selectivity to social vocalizations at the level of the midbrain. Furthermore, there is strong evidence that the IC is the first site of spectral integration in the auditory system (Marsh et al. 2006; Portfors and Wenstrup 2001), making the IC an ideal candidate for the first location in the ascending auditory system where selectivity to social vocalizations occurs.

Indeed, neural selectivity to social vocalizations has been found in the IC of two bat species (Klug et al. 2002; Portfors 2004). Moreover, there is evidence that multiply tuned and combination-sensitive neurons are involved in encoding social vocalizations in the mustached bat IC (Leroy and Wenstrup 2000; Portfors and Wenstrup 2002). For example, preliminary studies indicated that combination sensitivity is important for creating neural selectivity among social vocalizations (Portfors 2004). However, these analyses were qualitative in nature and therefore did not provide a thorough understanding of the neural mechanisms underlying encoding of social vocalizations in the IC.

The purpose of this study was to quantify the contribution of complex frequency interactions on the responses of IC neurons to social vocalizations. We hypothesized that the spectral integration characteristics (multiple tuning, facilitation, inhibition) of each IC neuron determine its responses to complex stimuli such as social vocalizations. Responses to synthetic single- and two-tone stimuli were recorded for a sample of neurons in the IC of the mustached bat. In addition, we obtained responses of the same neurons to social vocalizations of the mustached bat. We then designed three types of data-driven models of the stimulus-response relationship for each neuron using the tone stimuli and neural response data. All the models were then used to predict the responses of each neuron to the ensemble of social vocalizations. We show that the integration of multiple frequency bands has a significant effect

Address for reprint requests and other correspondence: C. V. Portfors, School of Biological Sciences, Washington State University, Vancouver, WA 98686 (E-mail: portfors@vancouver.wsu.edu).

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

on the responses of IC neurons to social vocalizations. This finding suggests that frequency integration properties (multiple tuning and combination sensitivity) are important neural mechanisms for discriminating among social vocalizations, and that these mechanisms are functioning at the level of the auditory midbrain.

METHODS

Electrophysiology

We recorded responses of well isolated single units in the IC of awake mustached bats (*Pteronotus parnellii*) to single tones, combinations of tones, and social vocalizations. Animals used in the experiments were captured from Trinidad and Tobago and maintained in a heated and humidified flight room at WSU Vancouver. The number of bats used was minimized by recording from an individual bat over multiple days. Animal care and use procedures were in strict accordance with the National Institutes of Health Guide for the Care and Use of Animals and were approved by the Washington State University Institutional Animal Care and Use Committee.

Surgical procedures

To enable recordings from single units in the IC of awake bats, we attached a metal pin onto the bat's skull and bolted this pin to a custom-made stereotaxic apparatus during electrophysiological recordings. The surgery to attach the pin was conducted 2 or 3 days prior to the beginning of electrophysiological recordings. The bat was anesthetized with isoflurane inhalation (IsoFlo, Abbott, North Chicago, IL) and placed in a custom-designed head holder with a bite bar to secure the head in a fixed position. Isoflurane inhalation was continued throughout the surgery via a tube placed over the bat's nostrils. To expose the skull, the skin and muscles were reflected laterally, and topical lidocaine was applied to the open wounds. The surface of the skull was cleared of tissue, and the pin was cemented onto the skull using ultraviolet-cured dental cement. A tungsten ground electrode was cemented into the right cerebral cortex. Using a scalpel, a small hole was made over the IC using landmarks visible through the skull. A local anesthetic (lidocaine) and a topical antibiotic (Neosporin) were applied to the wound and the animal was placed in a home cage. Once the surgery was performed, the bat was not returned to the main flight room. Electrophysiological recordings were started 2–3 days after the surgery.

Acoustic stimulation

Pure tones and natural mustached bat social vocalizations were presented as stimuli. The pure tone stimuli were synthesized using custom-written C++ computer algorithms. The social vocalizations consisted of a suite of 15 calls that are regularly emitted by mustached bats living in captivity (Kanwal et al. 1994). Sound stimuli were output through a high speed, 16-bit D/A converter (Microstar Labs; 400,000 samples/s), fed to a programmable attenuator (Tucker Davis Technologies PA5), a power amplifier (Parasound) and then to a leaf tweeter (Emit) speaker located 10 cm away from the bat and oriented at an angle of 25° toward the contralateral ear. The vocalizations were adjusted so that they were all output at the same peak intensity. The acoustic properties of the system were regularly tested using a 1/4-in calibrated microphone (Bruel and Kjaer, model 4135) placed in the position occupied by the bat's ear during experiments. There was a smooth, gradual decrease in sound pressure from 6 to 100 kHz of ~2.7 dB per 10 kHz. We measured the intensity of the second and third harmonic distortion components using custom-designed software performing a fast Fourier transform of the digitized microphone signal. Distortion components were buried in the noise floor, ≥ 50 dB below the signal floor.

Extracellular recording procedures and data acquisition

During electrophysiological recordings, the bat was placed in a foam restraining device that was molded to the bat's body. The bat's head was maintained in a uniform position by attaching the pin on the bat's head to a bar on a custom-designed stereotaxic apparatus. The stereotaxic apparatus was on an air table located in a single-walled sound-attenuating chamber that was heated and covered with acoustic foam. After the animal was fixed in place, the electrode was positioned over the IC while viewing with a surgical microscope. The electrode was advanced until it penetrated the IC and then the surface of the brain was covered with petroleum jelly to prevent drying. The electrode was subsequently advanced through the IC using a hydraulic micropositioner (David Kopf Instruments, Tujunga, CA) located outside the sound chamber. Acoustic stimulus generation and data acquisition were controlled outside the chamber by a computer running custom-written software.

Recording sessions typically lasted 4–5 h. During the recording session, the experimenter offered the bat water after each electrode penetration, which was typically 2–3 h. The bats typically remained quiet throughout the duration of the experiments. If a bat showed signs of discomfort during an experiment, a mild sedative (acepromazine, 1 mg/kg ip) was given. The acepromazine apparently had no effect on the response properties of IC neurons. If the bat continued to struggle, the experiment was terminated for the day. Bats were not recorded from on consecutive days. Between recording sessions, petroleum jelly and bone wax were used to protect the exposed brain from drying.

To obtain recordings of single units with high signal-to-noise ratios, we used micropipettes filled with 1 M NaCl (resistances of 20–30 M Ω). Extracellular action potentials were amplified (Dagan), filtered (band-pass, 500–6,000 Hz; Krohn-Hite), and sent through an analog spike enhancer (Fredrick Haer) before being digitized by a 16-bit A/D converter (Microstar Labs, 10 000 sample/s). Individual neuronal waveforms, raster plots, peristimulus time histograms (PSTHs), and response statistics were displayed on-line and stored for off-line analyses using custom-written software. All data were output for further data analyses using custom routines written in Matlab (MathWorks, Natick, MA) and IGORPro (WaveMetrics, Lake Oswego, OR) software.

We made dorsal to ventral penetrations through the IC with different angles caudal to rostral and lateral to medial to record single units with best frequencies between 10 and 80 kHz. Because of the known tonotopy of the mustached bat IC, we were able to use appropriate pure tone stimuli as search stimuli. Once a single unit was well isolated, the best frequency (BF) and minimum threshold of the unit was obtained audiovisually. These values were later confirmed by the modeling procedures (described in the following text). For the purpose of collecting the neurophysiological data, we defined BF as the frequency requiring the lowest intensity to elicit stimulus-locked spikes to 50% of the presentations, and minimum threshold as the lowest intensity required to evoke a response to 50% of the stimuli at the BF.

To obtain excitatory frequency tuning curves, we presented single pure tone stimuli (10- to 50-ms duration, 0.5-ms rise/fall time, 4/s) of varying frequency and intensity. The frequencies of the tones varied between 10 and 100 kHz in 1- to 2-kHz steps. The intensity of each tone varied in 10- or 20-dB steps starting at 10 dB above threshold. Each frequency and intensity pair was presented 5–10 times. Spike counts were collected in a 200-ms recording window during each stimulus presentation. The neuron's rate of spontaneous activity was obtained by counting spikes within a 200-ms window when no stimulus was present.

To obtain facilitatory and inhibitory frequency response maps, we presented combinations of tones using a two-tone paradigm. For each tone pair, one tone was set at BF and at an intensity 10 dB above threshold. The frequency and intensity of the BF tone did not change throughout the two-tone testing paradigm. A second tone, presented

simultaneously (same duration as BF tone, 0.5-ms rise/fall time), was varied in frequency between 10 and 100 kHz in steps of 1–2 kHz and in intensity in 10- to 20-dB steps starting at 80 dB SPL and ending at the threshold of BF. Each two-tone combination was presented 5–10 times, and spike counts were obtained in a 200-ms recording window. Frequencies that suppressed the BF excitatory response by 20% comprised the inhibitory tuning curve of that neuron. Frequencies that facilitated the excitatory response (20% increase in response over the linear sum of the 2 frequencies) comprised the facilitatory tuning curve of that neuron. For some neurons, responses were only obtained using the two-tone stimuli with the intensity of the second tone set at 20 dB attenuation (~80 dB SPL). If there was no apparent two-tone facilitation or inhibition when the second tone had this high of an intensity, we often discontinued the two-tone test and started collecting responses to social vocalizations to maximize our likelihood of holding the single unit long enough to present all the social vocalization stimuli.

Neuronal responses to social vocalizations were then obtained by presenting the suite of 15 vocalizations (variable duration, 0.5-ms rise/fall time, 4/s, 200-ms recording window) 10–20 times at multiple intensities. Vocalizations were presented at a minimum of 40, 60, and 80 dB SPL based on the normal intensity that these calls are emitted. If a neuron responded with sound-evoked spikes to any of the vocalizations, then those vocalizations were presented in steps of 10 dB from 80 dB SPL down to the neuron's threshold at BF. In some units, variants of each vocalization were presented. These variants had slightly shifted frequencies that were 1, 2, or 3 SD above or below the mean frequency of that vocalization emitted by a number of different bats (Kanwal et al. 1994). The purpose of presenting variants of the same social vocalization was to encompass the variety in frequency that occurs with naturally occurring vocalizations emitted from different animals (Kanwal et al. 1994). In most cases, we found that if a single unit responded to a particular social vocalization, it responded to all the variants as well. Because of this finding, in later experiments, we only presented one social vocalization of each type (each neuron was presented with ≥ 15 different vocalizations).

STATISTICAL DETERMINATION OF EXCITATORY AND INHIBITORY FREQUENCY TUNING CURVES. For single-tone tests, multiple presentations were made for each frequency/intensity combination, and multiple recordings of equal duration were also made in the absence of any stimulus to measure the spontaneous spike rate of each neuron. For each pure tone frequency/intensity pair, we calculated the mean spike rate and its variance and compared these to the spontaneous activity using a two-tailed *t*-test with a significance level of $\alpha = 0.01$. If the tone response was significantly greater than the spontaneous activity, that frequency/intensity pair was classified as excitatory. If the tone response was significantly less than the spontaneous activity, that frequency/intensity was classified as inhibitory. To determine inhibitory tuning curves if there was no spontaneous activity, the same statistical approach was used except that the responses from each frequency/intensity pair were compared with a baseline response generated by presenting a BF tone at 10 dB SPL above threshold (2-tone test). If the response from the two tones was significantly lower than the BF baseline response, that frequency/intensity pair was classified as inhibitory. Because no responses were classified as facilitatory during electrophysiological recordings, we did not perform a statistical test to quantify facilitation.

Modeling the acoustic stimulus/neural response relationship for each neuron

To quantify the contribution of complex excitatory and inhibitory frequency interactions on the responses of neurons to social vocalizations, we developed three types of data-driven models of the stimulus-response relationship for each neuron using synthetic tone stimuli and spike response data. The first model type (*primary tuning model*) was

based on the classical view of tonotopic organization in the IC, i.e., each neuron has one excitatory tuning curve centered around its BF. The second model type (*extended tuning model*) was developed to account for recent findings that neurons in the IC have multiple excitatory and inhibitory frequency tuning curves (Portfors and Wenstrup 2002). The third model type (*2-tone model*) was constructed to account for facilitatory tuning curves or inhibitory tuning curves when the neuron had no spontaneous activity. All types of models were used to predict the responses of each neuron to the ensemble of social vocalizations. The predictions for each model were then compared with the actual neural responses by calculating a mean squared error. Differences between the mean squared errors from the three model types provided a means for quantifying the contribution of complex frequency tuning in determining responses to social vocalization in the IC.

For each neuron in the experiment ($n = 50$), the stimulus-response models were optimized to reproduce the relationship between the one- and two-tone stimulus protocols and the resulting firing rate of the neuron, as approximated by the PSTH. We designed the stimulus-response model as a discrete (in both frequency and time) linear finite impulse response (FIR) filter h_i such that

$$\hat{r}(t) = c + \sum_{i=1}^{n_f} \sum_{j=1}^{n_t} h_i(j) s_i(t-j) \quad (1)$$

where $\hat{r}(t)$ is the predicted time-varying firing rate, c is the spontaneous firing rate of the neuron, i is the frequency band index of the stimulus, j is the time lag index, $s_i(t)$ is the discrete spectrographic representation of the time-varying stimulus, n_f is the number of frequency bands, and n_t is the number of time lag indices. The spontaneous rate was included as a model parameter so that the model accurately predicted the neural activity in the absence of a stimulus.

MODEL OPTIMIZATION PROCEDURE. A two-step procedure was used to optimize the parameters of the FIR filter, h_i . First, a linear minimum mean square error (LMMSE) algorithm was used to find the best linear parameter that minimized the mean-square error between the model responses and the physiological responses to the tone stimuli. Second, the results of the first step were further optimized using a gradient descent algorithm on a nonlinear optimality criterion that nullified negative spike-rates in the model predictions. Each model was tuned on the full set of tone stimuli (all frequencies and intensities) that were presented to each neuron.

STEP 1, LMMSE PARAMETER DETERMINATION. Optimal values of h_i and c were uniquely determined for each model for the minimum mean square error by solving the *normal equations* (Manolakis et al. 2005) that arise from linear optimization theory. For each time step t of each stimulus presentation, the spectrogram of the stimulus signal was calculated from the current time step back to the maximum time lag, resulting in an $n_f \times n_t$ time-frequency (spectrographic) representation of the stimulus. Consecutive $n_f \times n_t$ blocks extracted in this fashion overlapped in $n_t - 1$ columns. This spectrogram was reshaped to form a single $n_f \cdot n_t$ length row of the input matrix \mathbf{S} . For each time step t , the neuron's normalized PSTH at that time became an entry in the output column vector \mathbf{R} .

From the stimulus and response matrices \mathbf{S} and \mathbf{R} , we calculated the time averaged correlation matrix \mathbf{C}_{ss} of \mathbf{S} and the cross-correlation vector \mathbf{C}_{sr} of \mathbf{S} and \mathbf{R} . The relationship between \mathbf{C}_{ss} , \mathbf{C}_{sr} , and the optimal LMMSE model parameters, \mathbf{h} , was expressed in the form of the normal equations: $\mathbf{C}_{ss} \mathbf{h} = \mathbf{C}_{sr}$. The unique optimal parameter vector, \mathbf{h} , which encompasses both h_i and c from Eq. 1 was then solved for by inverting the time averaged correlation matrix: $\mathbf{h} = \mathbf{C}_{ss}^{-1} \mathbf{C}_{sr}$.

STEP 2, NONLINEAR PARAMETER FITTING. Using the LMMSE solution as a starting point, a gradient descent algorithm was employed to further refine the parameters of the model for each neuron. The

nonlinearity in the error function forced negative predicted rates to be equal to zero

$$E_{non-linear} = \frac{1}{\sigma_r^2} \sum_i (r(t) - H(\hat{r}(t)))^2 \quad (2)$$

where $H(\hat{r})$ is the Heaviside function. Because the recorded response of a neuron is nonnegative but the model can predict a negative response, this nonlinearity (in the optimality criteria, not the model) was introduced to focus the fitting of the model on the nonnegative response of the neuron without penalizing it for negative predicted firing rates.

For each training epoch (1 pass through the input/output data set) the gradient of the error surface (defined by Eq. 2) with respect to each model parameter was determined and then updated iteratively in the direction of this gradient to reduce the error with respect to the optimality criteria. For each model, 5,000 training epochs were performed while the learning rate was annealed from 0.001 to 0.0001. The nonlinear error criteria and gradient descent method provided a much better fit to the PSTH of the actual response as demonstrated for one neuron in Fig. 1. Therefore we used the nonlinear error criteria in all of the models.

Primary tuning models were based on the classical view of tonotopic organization in the IC, i.e., each neuron has one excitatory frequency tuning curve centered around its BF. For the modeling procedure, the BF was defined as the frequency of the pure tone stimulus that elicited a statistically significant excitatory response at the lowest stimulus intensity presented. This intensity was always 10 dB above the neuron's actual threshold because during electrophysiological recordings we did not obtain responses to tones at threshold but started the tests at 10 dB above threshold. However, in comparing the BFs generated during neurophysiological data collection and from the model, we found no difference. It should be noted that in all plots of frequency tuning curves, the tuning curves do not contain BF or threshold but start at 10 dB above threshold. The frequency range of the primary tuning curve was defined to be ≥ 20 kHz wide and to terminate ≥ 5 kHz into a region surrounding the BF that did not respond significantly to pure tone stimuli. The primary tuning models had frequency bands equal to the number of unique frequency tones used in the training of the model that fell within the primary tuning range for the neuron. For example, Fig. 2A shows a primary tuning model for a neuron with a primary tuning curve ranging from 50 to 70 kHz.

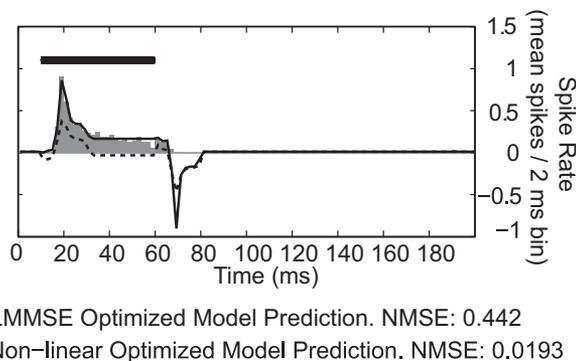


FIG. 1. Constructing the stimulus-response model. Neural response data [gray peristimulus time histograms (PSTH)] shows spike rate response to a tone stimulus. Horizontal bar represents stimulus. The model predictions are shown by the dashed and solid lines. The model was first optimized using an linear minimum square error (LMMSE) algorithm (dashed line) and then refined using a gradient descent algorithm after introducing a nonlinearity in the optimality criteria. The nonlinearly fit model (solid line) provided a much better prediction of the pure tone response than did the LMMSE optimized model.

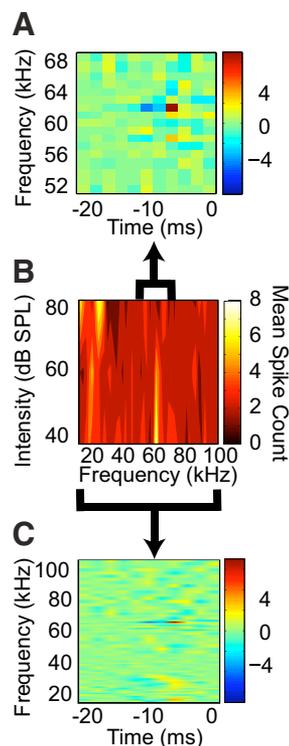


FIG. 2. Primary and extended model types were based on frequency tuning curves. *A*: primary tuning model was generated using frequencies and intensities only near the best frequency. *B*: frequency tuning curve of the neuron was used to determine the frequency range for the primary tuning model and to generate the extended model. *C*: extended model was generated using all frequencies and intensities in the single-tone stimulus protocol.

Extended tuning models contained inputs from the full frequency range used in the stimulus protocol (10–100 kHz). The extended tuning models had frequency bands equal to the number of unique frequency tones used in the training of the model. Figure 2C shows an extended tuning model for the neuron whose primary tuning curve is illustrated in A.

TWO-TONE MODELS. For the neurons where a two-tone stimulus protocol was applied, the same model optimization procedure was used except that both the single and two-tone data were used as the training set.

MODEL PREDICTIONS. Once the model types were generated for each neuron from the pure tone stimuli, they were used to make predictions of the neural responses to complex social vocalizations. The vocalizations were preprocessed to yield two spectrographic representations; the first was constrained to the frequency domain of the primary tuning model, and the second was unconstrained (0–100 kHz) for use with the extended model. These spectrographic representations were then convolved in the time domain with the corresponding models to generate the predicted, time-varying firing rate of the neuron.

INTERPRETATION OF RESULTS. The responses predicted by the three types of models were compared by a normalized mean-square error metric to quantify the differences in the predicted spike responses between the models. In addition, the difference between the FIR kernels for each model type was used to quantify the frequency interactions that were present only in the extended and two-tone models. Both types of model comparisons, spike response and FIR kernels, showed whether broad frequency interactions were present.

RESULTS

We recorded responses of 50 single units in the IC of awake mustached bats to single tones, combinations of tones, and social vocalizations. The BFs of these neurons ranged from 11 to 91 kHz. Using the statistical measure described in METHODS, each neuron was tested for excitatory, facilitatory, and inhibitory secondary tuning curves based on the results of the single- and two-tone tests. Thirteen of the neurons were singly tuned, whereas the remaining 37 were multiply tuned. All of the multiply tuned neurons had excitatory secondary tuning curves. Twenty-four of the 50 units had inhibitory secondary tuning curves that were more than one octave from BF. None of the neurons had facilitatory frequency tuning curves.

One of the key findings of this study in terms of understanding how single neurons in the IC respond to complex social vocalizations was that some neurons with BFs of ~60 kHz also had a secondary tuning curve in the 10- to 30-kHz range. Eleven of the 50 neurons had these types of tuning curves. Figure 3 shows the tuning curves for 8 of these types of neurons. In all cases, the primary tuning curves were sharp with BFs around 60 kHz, and the secondary tuning curves were broad with best responses in the 10- to 30-kHz range.

Model prediction accuracy

Of the 50 neurons used in this study, 42 had at least one response to a complex vocalization that was well modeled by the pure tone driven linear FIR models. Forty percent of the total responses to social vocalizations were modeled with NMSE values of <1 by either the neuron's primary or extended tuning models. Four of the 50 models were considered to be poor fits to the vocalization data and not useful for further analysis. The classification of "poor fit" was used because these models did not produce predictions to any of the vocalizations that were good in either the NMSE sense or in the sense that they captured the salient qualities of the neural responses. Interestingly, all four of these models performed reasonably well on the pure tone training stimuli, perhaps indicating extreme nonlinearities in the response characteristics

of some IC neurons when presented with complex stimuli. An additional four neurons had negligible responses to the vocalization stimuli and were not further considered in the modeling component of this study.

As indicated in Eq. 2, the NMSE has been normalized to the variance of each PSTH, meaning that NMSE of <1 indicates that the model prediction accurately accounted for some of the neural response. The mean and median NMSE values over all of the 42 models trained on the single-tone protocol were 8.8 and 1.1 for the primary tuning models and 10.5 and 1.6 for the extended models, respectively. The mean and median NMSE values over the 26 neurons for which the two-tone protocol was used were 5.9 and 1.2, respectively. There are three primary causes for the large mean error values.

First, for each neuron there was typically a large subset of the predictions that were accurate in the NMSE sense (NMSE < 1) along with a smaller number of outliers for which the predictions were poor. This was particularly true for the extended models, which tended to provide some of the best predictions along with some of the worst predictions due partially to the larger number of parameters needed to define the model. In addition, the NMSE calculation is particularly sensitive when very few spikes occur and results in high NMSE values.

Second, due to the relatively coarse time sampling of the stimulus and response in the predictions, small timing errors became large errors in the NMSE values even though the salient qualities of the neural response may have been captured well. For example, in Fig. 4, the model's prediction captured most of the amplitude and temporal qualities of the PSTH envelope even though the prediction NMSE value is 1.55.

Third, it is well known that neurons in the IC may integrate spectrally complex stimuli nonlinearly (Mittmann and Wenstrup 1995; Portfors and Felix 2005; Portfors and Wenstrup 1999). This is perhaps why in some neurons certain vocalization responses were not well predicted by the linear model. The finding that many of the responses to social vocalizations for 42 of the 50 neurons were suitably predicted by linear models indicates that neurons in the IC may often combine spectral

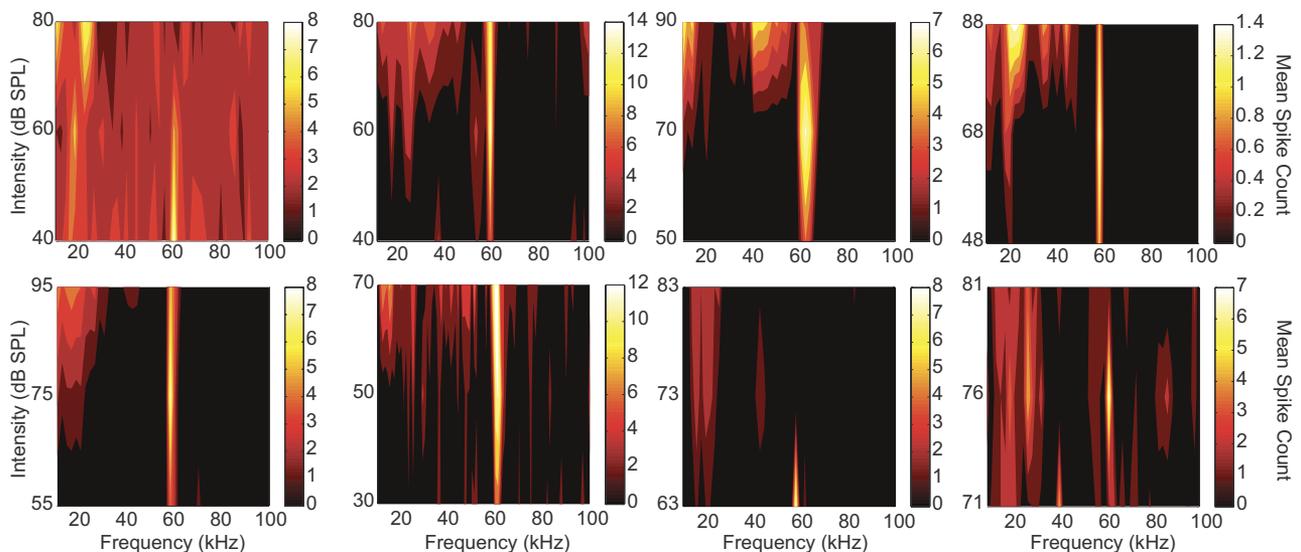


FIG. 3. Tuning curves from 8 of the neurons in the study that were characterized by a broad 10- to 30-kHz tuning and a narrow primary tuning at 60 kHz. These tuning curves do not show best frequency and threshold of the neurons but start at an intensity level 10 dB above threshold.

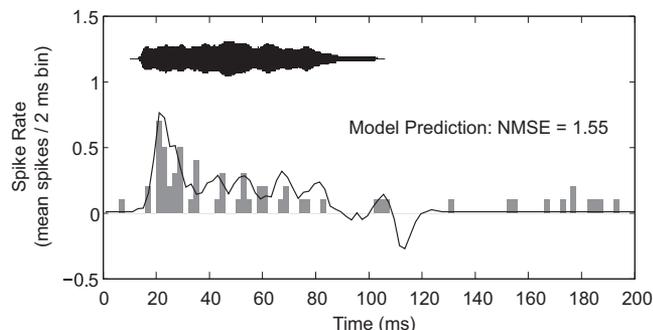


FIG. 4. Example of a model output that yielded poor predictions in the normalized mean squared error (NMSE) sense but still captured the salient qualities of the neuron's response. Time-amplitude plot of the vocalization stimulus is shown at the top of the figure. ■, neural response (PSTH) and —, model's prediction.

information in an approximately linear fashion. On the other hand, responses to social vocalizations with more pure tone structures tended to be better predicted by our models than responses to social vocalizations with broad band components perhaps suggesting that nonlinear frequency interactions are important in determining neural response to complex stimuli.

For the purpose of quantifying our results, we have chosen to define a "good" fit as one in which the vocalization prediction NMSE value was < 1 . Forty percent of the neural responses to social vocalizations in this study were modeled with

a NMSE value of < 1 by either the neuron's primary or extended tuning model.

Excitatory primary tuning curve contributions to social vocalizations

Nearly all of the neurons (40 of 42) had at least one response to a vocalization that was well modeled (NMSE < 1) by the neuron's primary tuning model. On average, the primary tuning models for these 40 neurons provided good (NMSE < 1) predictions for 37% of the vocalizations. In some neurons, the primary tuning models generated good predictions for all of the vocalizations that had energy in the neuron's primary tuning curve. Figure 5 shows the primary tuning model predictions for a neuron with responses to social vocalizations that were well predicted by the primary tuning model.

The tuning curve for this singly tuned neuron is illustrated in Fig. 5A along with the corresponding primary tuning model illustrated in B. Even though the tuning curve for this neuron indicates a slight response to frequencies close to 55 kHz (Fig. 5A), this neuron is classified as singly tuned because these responses were not found to be statistically significant. In this neuron, the primary tuning model encompassed the frequency range of 10–30 kHz. Predictions of responses to three social vocalizations based on this model are shown in Fig. 5C, 1–3. These three vocalizations contain the majority of their energy

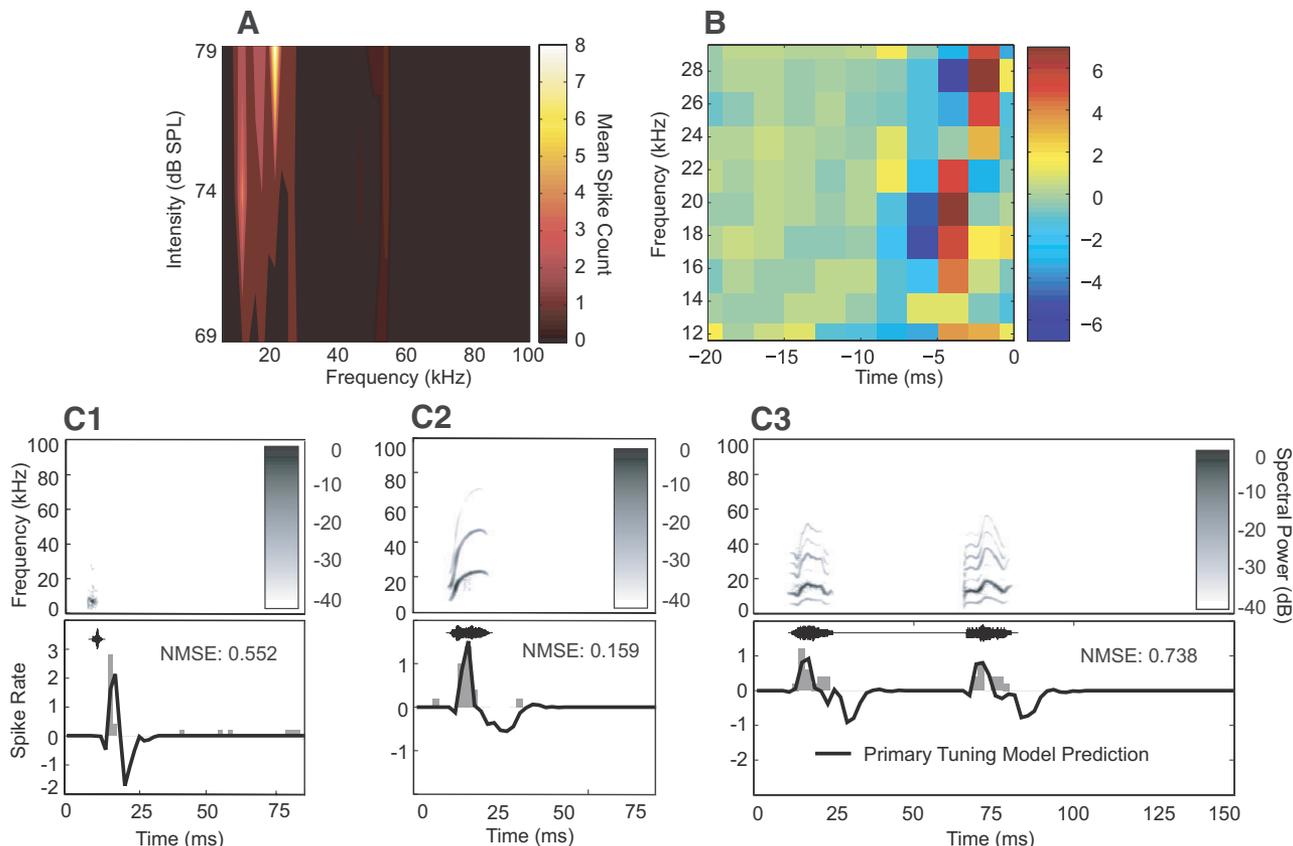


FIG. 5. Responses to social vocalizations of some neurons were modeled accurately by their primary tuning models. A: frequency tuning curve of the neuron. These tuning curves do not show best frequency and threshold of the neurons but start at an intensity level 10 dB above threshold. B: primary tuning model for the neuron. The primary tuning model encompassed a frequency range from 10 to 30 kHz. C, 1–3: predictions of the neuron's responses to 3 vocalizations along with the neuron's actual responses. Top: spectrograms of the vocalizations. Spectral power is the RMS sound pressure per Hz. The scale is normalized to the maximum energy in each vocalization. Bottom: time-amplitude plot of the vocalization stimulus, the neuron's response (gray PSTH), and the model's prediction (solid black line).

in the low-frequency range encompassing the primary tuning curve in the model. Consequently, the model provided good predictions of the responses with NMSE values of 0.552, 0.159, and 0.738. In all three cases, the primary tuning model predictions captured the salient features of the neural response. The extended or two-tone models in this example did not decrease the NMSE values, indicating that the primary tuning curve drove the responses to the vocalizations.

Excitatory secondary tuning curve contributions to responses to social vocalizations

Nearly all of the neurons (39 of 42) had at least one response to a vocalization that was well modeled (NMSE < 1) by the neuron's extended tuning model. On average, these 39 neurons provided good (NMSE < 1) predictions for 26% of the vocalizations. Even though this is a smaller fraction of the total predictions that match our "good fit" criteria than the primary tuning model predictions, the mean NMSE value for these extended model predictions was 0.73, whereas the mean NMSE value for the primary tuning model predictions on the same vocalizations was 0.88. For the cases where the extended models quantifiably "captured" more of the true neural response characteristics than the primary tuning models, we can compare the two model's predictions to determine causes of this behavior.

Figure 6 shows the results for a neuron where 65% of the vocalization responses were well predicted (NMSE < 1) by the neuron's extended model. Figure 6A shows the tuning curve for the neuron. This neuron had sharp tuning in the 60-kHz range as well as broad tuning in the 10- to 30-kHz range. As shown in Fig. 6B, the model predicted responses to pure tone stimuli in the primary (61 kHz—the neuron's BF) and secondary (20 kHz) frequency tuning ranges of the neuron very accurately. Figure 6C shows the primary tuning model, which spans between 50 and 68 kHz and illustrates a strong response to a 61-kHz stimulus. Figure 6D shows the extended model for the neuron, indicating a second strong excitatory response in the 10- to 30-kHz range as well as some excitatory response around 45 and 85 kHz. The rest of the plots (Fig. 6E, 1–6) show the neuron's responses to six different vocalizations along with the extended and primary tuning model predictions. Even though the BF of this neuron was 61 kHz, it responded strongly to nearly all of the vocalizations that had spectral power in the 10- to 30-kHz range. Of particular interest is how the primary tuning model predicted little or no response to this suite of vocalizations, whereas the extended model accurately predicted the neural responses. It is clear that this neuron responded to the vocalizations because of its broad 10- to 30-kHz tuning and not because of its primary frequency tuning curve. In this example, the NMSE values for the extended model were significantly lower than the NMSE values generated with the primary tuning model (*t*-test; $P \leq 0.001$).

Inhibitory secondary tuning curve contributions to responses to social vocalizations

We did not find any neurons where two-tone facilitation influenced how the neuron responded to vocalizations, but our neurophysiological results indicated that 24 of the 50 neurons had inhibitory secondary tunings. These inhibitory tuning

curves were an octave or more from BF and thus were not sideband or lateral inhibition. The inhibitory tuning curves were sometimes lower in frequency than the BF tuning curve and sometimes higher. Although not as evident as the excitatory secondary tuning curves, our modeling approach suggests that inhibitory tuning curves influence neural responses to vocalization stimuli. Figure 7 shows a neuron that had a secondary inhibitory tuning curve that influenced the way the neuron responded to the social vocalizations. Figure 7A shows the tuning curve for this neuron. This neuron had a BF of 47 kHz and was sharply tuned. It also had a secondary excitatory tuning curve between 10 and 20 kHz. Because the spontaneous firing rate for this neuron was near zero, the tuning curve does not show any inhibition. To determine inhibitory secondary tuning curves, we conducted two-tone tests where two tones were presented simultaneously. In the example illustrated in Fig. 7, one tone was always presented at BF at 10 dB above threshold (47 kHz at 40 dB SPL) and the second tone ranged from 10 to 100 kHz. The BF tone provided a baseline level of activity in the neuron so that inhibition as a result of the second tone could be measured. Figure 7B shows a cross-section of the two-tone tuning curve where a pure tone presented at 60 dB SPL ranged from 10–100 kHz. The *y* axis shows the difference in the mean spike count relative to the baseline activity (a value of 0 indicates the 2nd tone had no effect on the baseline response of the neuron). Here, it is clear that pure tones presented at 20 and 72 kHz strongly inhibited the baseline response of the neuron, whereas there was a low level inhibitory effect for all frequencies except those around the BF of the neuron.

Figure 7C shows the model fit from the single-tone tests for the neuron. Figure 7D shows the model fit from the combination of the single- and two-tone tests for the neuron. Figure 7E shows the difference between the two models. Note how fitting the model off of the two-tone tests increased the inhibition present in the model at 20 and 72 kHz.

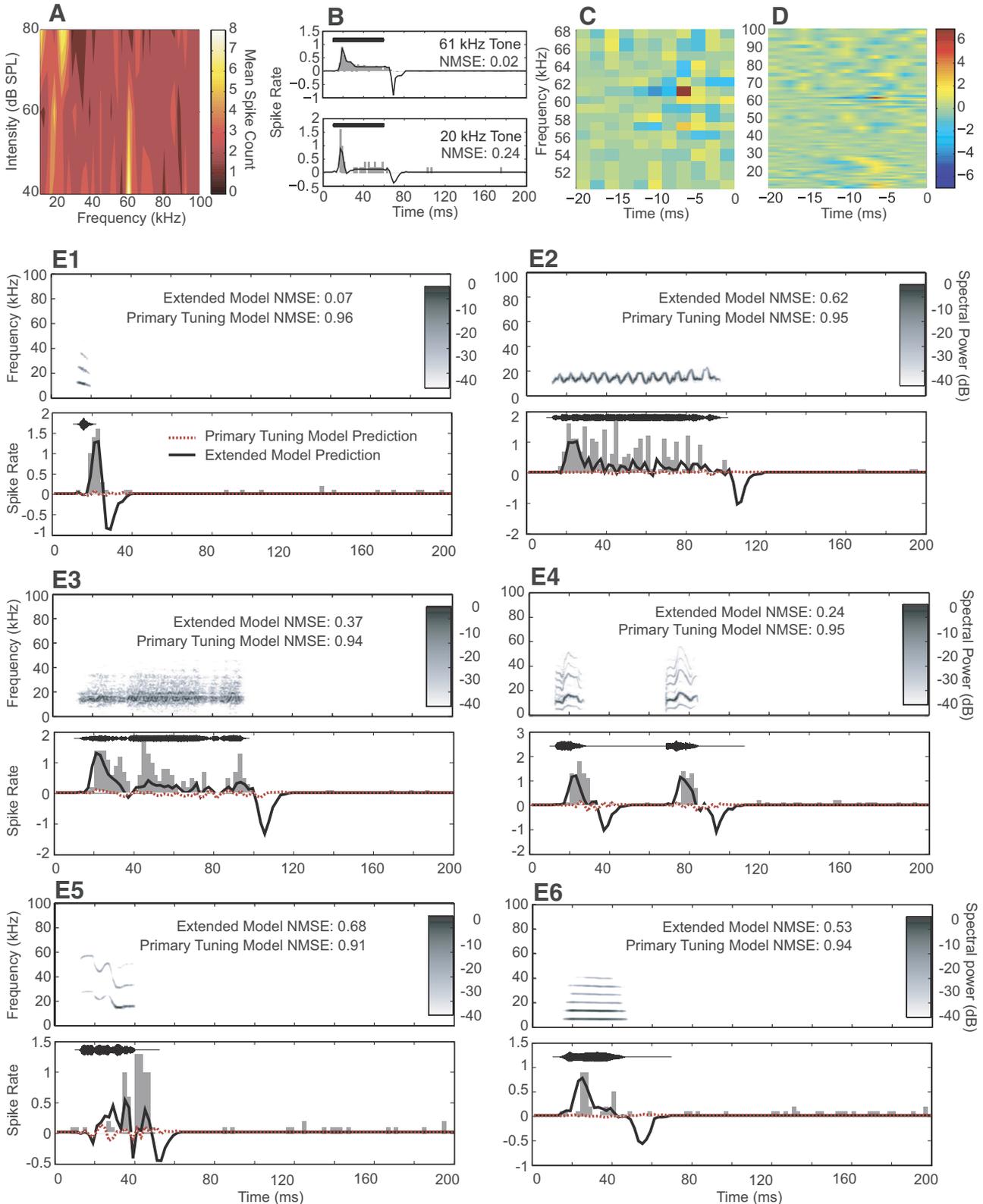
Incorporating the two-tone tests into our models allowed us to quantify the effect of inhibitory tuning curves on responses to vocalizations. Figure 7F, 1–4, shows that for three of the vocalizations, the predictions made by the two-tone model were substantially better than the one-tone model for this neuron. In Fig. 7F1, the predictions made by the one- and two-tone models are very similar; both predicted a response and both captured the salient features of the neuron's response to this particular vocalization. In Fig. 7F, 2–4, the two-tone model generated substantially better predictions with lower NMSE values. In all these examples, the social vocalizations contain spectral power in the 20-kHz range; the frequency range of one of the inhibitory tuning curves. In each example, the single-tone model predicted a strong response because there was also spectral power in the primary or secondary excitatory tuning curves of the neuron. In contrast, the two-tone model correctly predicted that the neuron would not respond because of the inhibitory tuning curve captured by this model in the 20-kHz range.

DISCUSSION

A major finding of this study is that many responses of neurons in the IC to complex social vocalizations are not well predicted by their primary excitatory tuning curve. Of the

responses that were well predicted (NMSE < 1) by the extended models in this study, the primary tuning models had 21% higher NMSE values (0.73 vs. 0.88). Furthermore, as exemplified in Fig. 6, some neurons had responses to vocalizations that could only be predicted by modeling responses outside of the neuron's primary tuning range. This suggests

that spectral integration features of neurons in the IC such as multiple tuning and combination-sensitive inhibition play a role in generating responses to social vocalizations. These results support the findings that in humans, speech can be recognized when frequency is degraded but is enhanced by selective spectral information (Shannon et al. 1995). Temporal



cues are also likely important for encoding social vocalizations in the mustached bat IC as has been shown for speech recognition (Shannon et al. 1995), but this remains to be tested in future studies. The modeling methods used in this study could be modified to quantify the contributions of temporal modulations in eliciting responses to complex stimuli in the IC.

By modeling the acoustic stimulus/neural response relationship for each neuron, we were able to quantify the contribution of secondary excitatory and inhibitory frequency tuning curves by comparing predictions made by the primary tuning model and the extended model. This advances our understanding of the contribution of spectral integration features of neurons when responding to complex sounds.

The modeling approach used in this study drew heavily from the spectro-temporal receptive field (STRF) methodologies (Aertsen and Johannesma 1981; Eggermont 1993; Eggermont et al. 1983) that have been used to model neurons throughout all levels of the ascending auditory pathway from the auditory nerve (Young and Calhoun 2005) to the auditory forebrain (Sen et al. 2001). We have included several modifications and extensions to common STRF implementations, however, that have aided us in the analysis presented in this study.

The common bond that ties STRF methodologies together is that they embody a linear mapping between spectro-temporal features of a stimulus and the resulting neural response. For input into the STRF, the stimuli are decomposed into frequency components to approximate the frequency decomposition performed by the cochlea and the tonotopic organization of auditory nuclei. Many approaches have been taken for the spectro-temporal representation of auditory stimuli, including spectrographs, wavelet and gamma tone transformations, and representations with adaptive gain control to more closely model the cochlear response (Gill et al. 2006; Slaney and Lyon 1993).

In our study, we chose a spectrographic representation in units of decibels of sound pressure level per hertz (dB SPL/Hz). This scaling was used to represent the stimulus in terms of spectral power on a logarithmic scale and to have a value of 0 at or near the threshold of hearing. No additional cochlear dynamics were included in this representation due to the absence of an appropriate cochlear model for the mustached bat, which is known to have an acoustic fovea at around 60 kHz.

Another decision that affects the fitting of the STRF model is the choice of stimuli. Historically, broad band Gaussian noise has been used to simplify the analysis of reverse correlation and spike triggered average techniques (Aertsen and Johannesma 1981). Broad band noise has limited use as a stimulus protocol, however, because many auditory neurons do not respond robustly to this stimulus (Theunissen et al. 2004)

Pure tones, narrow band Gaussian noise, ripple noise, and natural stimuli such as recorded vocalizations have also been explored (Klein et al. 2000; Theunissen et al. 2000). In our study, we were specifically interested in measuring how neural responses to complex stimuli such as social vocalizations were dependent on input from distinct frequency bands in the stimulus domain. Because of this, we chose to use only pure tone stimuli in the fitting of our models and made the simplifying assumption that the responses to complex stimuli are the linear summations of the pure tone responses. We are aware that this is an incomplete picture of the actual stimuli/response relationship for complex stimuli, but this approach allowed us to address the fundamental questions posed in this study. Using this linear approach, we obtained predictions with NMSE values of <1 in 40% of the total response predictions. These results suggest that many neurons in the IC combine spectral information in an approximately linear fashion, at least for some of their responses.

A final important model design decision concerns the methods used to fit the model and to evaluate its performance. Spike-triggered averaging techniques break down when statistical correlations exist in the stimuli. This is particularly true in the cases of pure tone stimuli and vocalization stimuli because both are highly structured. Methods have been developed that address this issue to allow for optimal model fitting even when using correlated stimuli (Theunissen et al. 2000). The first step of the fitting approach used in our study is related to these techniques and is further drawn from the field of optimal linear minimum mean squared error (LMMSE) signal modeling. A key distinction between our method and the methods common in the STRF literature is our use of a second step in the fitting process in which a nonlinearity is introduced in the optimality criteria (note: the model is still linear in the stimulus-response relationship, only the fitting procedure is nonlinear). Because the measured firing rate of the neuron can never be negative, this nonlinearity was introduced so as not to “penalize” the model when negative firing rates were predicted. Using the LMMSE fit as a starting point, a gradient descent method was employed to further refine the model with respect to this new optimality criteria. We found that this extension to the classical STRF approach greatly improved the predictive performance of our models, as measured by the NMSE between the actual and predicted firing rates of the neuron.

Although our models were based on linear summations, we do not discount the effect of nonlinear interactions in the IC and their influence on responses to complex stimuli. For example, the social vocalizations with more pure tone structures tended to be better predicted by our linear models compared with social vocalizations with more broad band

FIG. 6. Comparison of predicted responses to social vocalizations of the primary and extended tuning models for one neuron. *A*: frequency tuning curve for the neuron. Best frequency and threshold are not shown and tuning curve starts at an intensity of 10 dB above threshold. This neuron had sharp tuning in the 61-kHz range (the BF of the neuron) as well as broad tuning in the 10- to 30-kHz range. *B*: predicted pure tone responses for a 61-kHz tone (*top*) and a 20-kHz tone (*bottom*), showing a good NMSE fit of the model to the pure tone data. PSTH bin size is 2 ms. *C*: primary tuning model, which spanned 50–68 kHz and indicated a strong response to 61-kHz pure tone stimulus. *D*: extended model for the neuron, indicating additional strong excitation in the 10- to 30-kHz range as well as around 45 and 85 kHz. *E*, 1–6: comparison of primary and extended model predictions for the neuron’s response to 6 different social vocalizations. *Top*: spectrograms of the vocalizations. *Bottom*: time-amplitude plot of the vocalization stimulus, the neuron’s response (gray PSTH), the primary model’s prediction (dashed red line) and the extended model’s prediction (solid black line). Even though the BF of this neuron was 61 kHz, it responded strongly to nearly all of the vocalizations that had spectral power in the 10- to 30-kHz range. Of particular interest is how the primary tuning model predicted little or no response to this suite of vocalizations, while the extended model accurately predicted the neuron’s responses. It is clear that this neuron responded to the vocalizations due to its broad 10- to 30-kHz tuning and not because of its primary 61-kHz tuning. Normally this neuron would be considered as specialized for encoding echolocation signals, but here it responded vigorously to social vocalizations.

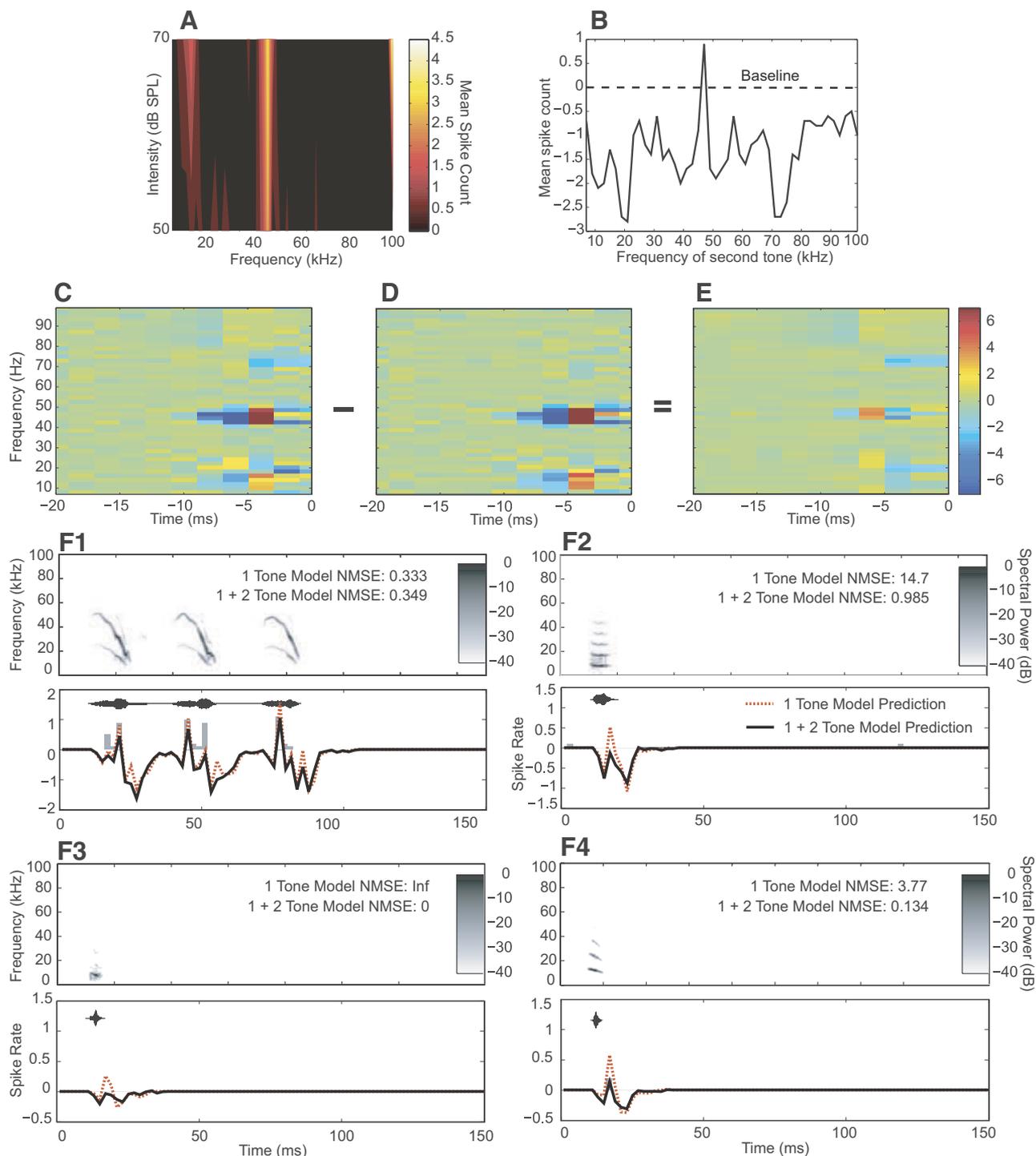


FIG. 7. Comparison of predicted responses to social vocalizations of the extended one-tone model and the 2-tone model. Inhibitory regions far from best frequency (BF) increased selectivity to specific social vocalizations. *A*: excitatory tuning curve for a single neuron. Best frequency and threshold are not shown. *B*: response of the neuron during a 2-tone test. One tone was set at the neuron's BF, 47 kHz, and the 2nd tone varied in frequency (x axis). The y axis shows the difference in the mean spike count relative to the baseline activity resulting from the BF tone; 0 is the baseline activity level of the BF tone. Tones presented at 20 and 72 kHz strongly inhibited the baseline response of the cell. *C*: extended model generated from the single- and 2-tone tests for the neuron. *E*: difference between the 2 models. Including the 2-tone tests in the training data increased the inhibition present in the model at 20 and 72 kHz. *F*, 1–4: comparison of the extended and 2-tone model predictions for the neuron's response to 6 different social vocalizations. *Top*: spectrograms of the vocalizations. *Bottom*: time-amplitude plot of the vocalization stimulus, the neuron's response (gray PSTH), the extended (1 tone) model's prediction (dashed red line) and the 2-tone model's prediction (solid black line). *F1*: response where both the single- and 2-tone models provided a similar prediction for the vocalization. This is because much of the power of this vocalization was within the primary tuning range of the neuron. *F*, 2–4: predictions to 3 vocalizations, the spectral power of which was predominantly in the 20-kHz range. In each of these, the single-tone model predicted a strong response, but the 2-tone model accurately predicted no response due to the inhibition captured by this model in the 20-kHz range.

components. In addition, we did not have any neurons in our sample that responded in a facilitatory combination-sensitive manner. These neurons show nonlinear interactions in that their responses to combinations of stimuli are greater than the linear sum of the responses to the stimuli (Mittmann and Wenstrup 1995; Portfors and Wenstrup 1999, 2002). Facilitatory combination-sensitive neurons have qualitatively been shown to be important in creating selective responses to social vocalizations in the IC (Portfors 2004). Thus it is of interest to further investigate the role of nonlinear interactions as mechanisms for processing complex sounds in the IC.

One of the key findings of this study was the high percentage of neurons in the mustached bat IC that had secondary excitatory frequency tuning curves and the effect these secondary tuning curves had on responses to social vocalizations. Almost 75% of the neurons had a distinctive secondary excitatory tuning curve that was often more than one octave away from the BF tuning curve. Furthermore, the modeling approach used in this study showed that 39 of 42 of the neurons had at least one response to a social vocalization that was better predicted when accounting for these secondary frequency tunings. A subset of the multiply tuned neurons was particularly striking in that their primary tuning was in the 60-kHz range; a region that is highly sensitive to echo information. These neurons had sharp tuning in the 60-kHz region and were also broadly tuned in the range of 10–30 kHz. Because many of the mustached bat social vocalizations contain energy within this low-frequency range (Kanwal et al. 1994), this low-frequency tuning caused the 60-kHz neurons to respond to many social vocalizations. This is a significant finding because the 60-kHz region of auditory nuclei in the mustached bat is hypertrophied and has always been thought to primarily be involved in encoding echo information (Olsen and Suga 1991; Suga 1989; Suga and Jen 1976; Suga et al. 1975). For instance, the sharp tuning of these neurons to 60 kHz makes them ideally suited to process echo information related to Doppler shift, which allows the bat to determine its relative velocity (Suga and Jen 1976; Suga et al. 1975). Our findings now suggest that these so-called echolocation neurons perform a second function; they encode social vocalizations.

Dual tasking of neurons has also been shown in the auditory cortex of mustached bats (Ohlemiller et al. 1996). In regions known to encode specific elements found in echolocation signals (FM-FM area), some neurons also respond to combinations of elements found in social vocalizations (Esser et al. 1997). In addition, dual tasking of neurons occurs in the auditory cortex of the pallid bat where individual neurons are tuned to both low and high frequencies and respond to noise transients used during prey detection and ultrasonic FM sweeps used during echolocation tasks (Razak et al. 1999). Neurons that are multiply tuned can thus change their response properties based on behavioral context. In bats, these examples of dual tasking relate to individual neurons being recruited for processing active (echolocation) or passive hearing (prey detection, communication). However, this does not mean that neurons with dual tasks are limited to specialized auditory systems. Multiply tuned neurons are also common in the auditory cortex and IC of less specialized mammals such as mice (Portfors and Felix 2005). A novel finding of our study is that neurons with dual functions occur in the IC. Previous

studies of dual task neurons have all been in the auditory cortex.

Another novel finding of this study is that inhibitory frequency tuning curves that are far away from the excitatory tuning curve (i.e., not sideband inhibition) have a strong influence on responses to vocalizations. In neurons with combination-sensitive inhibition, responses to social vocalizations were poorly predicted by our model that only incorporated the excitatory tuning curve. For vocalizations that included energy in the neuron's excitatory and inhibitory tuning curves, the single-tone model predicted a response. However, the energy in the inhibitory tuning region suppressed the neural response. The responses to vocalizations were then well predicted by including the inhibitory tuning curve into the model.

The finding that neurons in IC have inhibitory regions off of BF that influence their responses to vocalizations suggests a mechanism for creating selectivity to particular vocalizations. Inhibition around BF is known to increase selectivity to vocalizations (Klug et al. 2002), but our study suggests that many neurons also receive inhibitory projections that are tuned to frequencies often up to an octave away from BF. These inhibitory inputs can profoundly affect how a neuron responds to complex sounds.

The findings from this study suggest that a thorough understanding of how complex sounds are processed in auditory nuclei requires more than an understanding of the tonotopic gradient of the structure. In particular, in the IC we found that the majority (75%) of neurons were multiply tuned and our modeling approach showed that many of these neurons had responses to social vocalizations that were significantly influenced by the secondary (i.e., nontonotopic) excitatory and inhibitory frequency tunings. This means that stimulating the IC based on its tonotopic gradient alone, as has recently been done with auditory midbrain implants (Lenarz et al. 2006a,b; Samii et al. 2007), is not likely to be the most optimal method of restoring speech perception in patients with sensorineural hearing loss.

ACKNOWLEDGMENTS

We thank J. McNames for invaluable advice in developing the modeling framework. We thank two anonymous reviewers for providing input on this manuscript.

GRANTS

This work was supported by the National Institute on Deafness and Communication Disorders Grant R01 DC-04733 to C. V. Portfors and National Science Foundation Grants IOB-0445648 to P. D. Roberts and IOS-0620560 to C. V. Portfors.

REFERENCES

- Adams JC. Ascending projections to the inferior colliculus. *J Comp Neurol* 183: 519–538, 1979.
- Aertsen AM, Johannesma PI. The spectro-temporal receptive field: a functional characteristic of auditory neurons. *Biol Cybern* 42: 133–143, 1981.
- Eggermont JJ. Wiener and volterra analyses applied to the auditory system. *Hear Res* 66: 177–201, 1993.
- Eggermont JJ, Aertsen AM, Johannesma PI. Prediction of the responses of auditory neurons in the midbrain of the grass frog based on the spectro-temporal receptive field. *Hear Res* 10: 191–202, 1983.
- Esser KH, Condon CJ, Suga N, Kanwal JS. Syntax processing by auditory cortical neurons in the fm-fm area of the mustached bat *Pteronotus parnellii*. *Proc Natl Acad Sci USA* 94: 14019–14024, 1997.
- Gill P, Zhang J, Woolley SMN, Fremouw T, Theunissen FE. Sound representation methods for spectro-temporal receptive field estimation. *J Comput Neurosci* 21: 5–20, 2006.

- Kanwal JS, Matsumura S, Ohlemiller K, Suga N.** Analysis of acoustic elements and syntax in communication sounds emitted by mustached bats. *J Acoust Soc Am* 96: 1229–1254, 1994.
- Klein DJ, Depireux DA, Simon JZ, Shamma SA.** Robust spectrotemporal reverse correlation for the auditory system: optimizing stimulus design. *J Comput Neurosci* 9: 85–111, 2000.
- Klug A, Bauer EE, Hanson JT, Hurley L, Meitzen J, Pollak GD.** Response selectivity for species-specific calls in the inferior colliculus of mexican free-tailed bats is generated by inhibition. *J Neurophysiol* 88: 1941–1954, 2002.
- Lenarz M, Lim HH, Patrick JF, Anderson DJ, Lenarz T.** Electrophysiological validation of a human prototype auditory midbrain implant in a guinea pig model. *J Assoc Res Otolaryngol* 7: 383–398, 2006a.
- Lenarz T, Lim HH, Reuter G, Patrick JF, Lenarz M.** The auditory midbrain implant: a new auditory prosthesis for neural deafness-concept and device description. *Otol Neurotol* 27: 838–843, 2006b.
- Leroy SA, Wenstrup JJ.** Spectral integration in the inferior colliculus of the mustached bat. *J Neurosci* 20: 8533–8541, 2000.
- Lippe W, Rubel E.** Development of the place principle: tonotopic organization. *Science* 219: 514–516, 1983.
- Manolakis DG, Ingle VK, Kogon SM.** *Statistical and Adaptive Signal Processing*. Boston, MA: Artech House, 2005.
- Marsh RA, Nataraj K, Gans D, Portfors CV, Wenstrup JJ.** Auditory responses in the cochlear nucleus of awake mustached bats: precursors to spectral integration in the auditory midbrain *J Neurophysiol* 95: 88–105, 2006.
- Mittmann DH, Wenstrup JJ.** Combination-sensitive neurons in the inferior colliculus. *Hear Res* 90: 185–191, 1995.
- Ohlemiller K, Kanwal J, Suga N.** Facilitative responses to species-specific calls in cortical fm-fm neurons of the mustached bat. *Neuroreport* 7: 1749–1755, 1996.
- Olsen JF, Suga N.** Combination-sensitive neurons in the medial geniculate body of the mustached bat: encoding of relative velocity information. *J Neurophysiol* 65: 1254–1274, 1991.
- Portfors C.** Combination sensitivity and processing of communication calls in the inferior colliculus of the moustached bat *Pteronotus parnellii*. *An Acad Bras Cienc* 76: 253–257, 2004.
- Portfors C, Felix R 2nd.** Spectral integration in the inferior colliculus of the cba/caj mouse. *Neuroscience* 136: 1159–1170, 2005.
- Portfors CV, Wenstrup JJ.** Delay-tuned neurons in the inferior colliculus of the mustached bat: implications for analyses of target distance. *J Neurophysiol* 82: 1326–1338, 1999.
- Portfors CV, Wenstrup JJ.** Responses to combinations of tones in the nuclei of the lateral lemniscus. *J Assoc Res Otolaryngol* 2: 104–117, 2001.
- Portfors CV, Wenstrup JJ.** Excitatory and facilitatory frequency response areas in the inferior colliculus of the mustached bat. *Hear Res* 168: 131–138, 2002.
- Razak KA, Fuzessery ZM, Lohuis TD.** Single cortical neurons serve both echolocation and passive sound localization. *J Neurophysiol* 81: 1438–1442, 1999.
- Samii A, Lenarz M, Majdani O, Lim HH, Samii M, Lenarz T.** Auditory midbrain implant: a combined approach for vestibular schwannoma surgery and device implantation. *Otol Neurotol* 28: 31–38, 2007.
- Sen K, Theunissen FE, Doupe AJ.** Feature analysis of natural sounds in the songbird auditory forebrain. *J Neurophysiol* 86: 1445–1458, 2001.
- Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M.** Speech recognition with primarily temporal cues. *Science* 270: 303–304, 1995.
- Slaney M, Lyon RF.** On the importance of time - a temporal representation of sound. In: *Visual Representations of Speech Signals*, edited by Cooke M, Beet SW, Crawford M. New York: Wiley, 1993, p. 95–116.
- Suga N.** Principles of auditory information-processing derived from neuroethology. *J Exp Biol* 146: 277–286, 1989.
- Suga N, Jen PHS.** Disproportionate tonotopic representation for processing cf-fm sonar signals in the mustache bat auditory cortex. *Science* 194: 542–544, 1976.
- Suga N, Simmons JA, Jen PHS.** Peripheral specialization for fine analysis of doppler-shifted echoes in the auditory system of the “cf-fm” bat *Pteronotus parnellii*. *J Exp Biol* 63: 161–192, 1975.
- Theunissen FE, Sen K, Doupe AJ.** Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. *J Neurosci* 20: 2315–2331, 2000.
- Theunissen FE, Woolley SMN, Hsu A, Fremouw T.** Methods for the analysis of auditory processing in the brain. *Ann NY Acad Sci* 1016: 187–207, 2004.
- Young ED, Calhoun BM.** Nonlinear modeling of auditory-nerve rate responses to wideband stimuli. *J Neurophysiol* 94: 4441–4454, 2005.
- Zook J, Winer J, Pollak G, Bodenhamer R.** Topology of the central nucleus of the mustache bat’s inferior colliculus: correlation of single unit properties and neuronal architecture. *J Comp Neurol* 231: 530–546, 1985.