Published in final edited form as:

Hear Res. 2010 August ; 267(0): 96–110. doi:10.1016/j.heares.2010.04.005.

Patterned Tone Sequences Reveal Non-Linear Interactions in Auditory Spectrotemporal Receptive Fields in the Inferior Colliculus

W. Owen Brimijoin^{1,3} and William E. O'Neill^{1,2}

¹Department of Brain and Cognitive Sciences, College of Arts, Science, and Engineering, University of Rochester, Rochester, NY 14627

²Center for Navigation and Communication Sciences, Department of Neurobiology & Anatomy, University of Rochester School of Medicine and Dentistry, Rochester, NY 14642-8603 USA

Abstract

Linear measures of auditory receptive fields do not always fully account for a neuron's response to spectrotemporally-complex signals such as frequency-modulated sweeps (FM) and communication sounds. A possible source of this discrepancy is cross-frequency interactions, common response properties which may be missed by linear receptive fields but captured using two-tone masking. Using a patterned tonal sequence that included a balanced set of all possible tone-to-tone transitions, we have here combined the spectrotemporal receptive field with two-tone masking to measure *spectrotemporal response maps* (STRM). Recording from single units in the mustached bat inferior colliculus, we found significant non-linear interactions between sequential tones in all sampled units. In particular, tone-pair STRMs revealed three common features not visible in linear single-tone STRMs: 1) two-tone facilitative interactions, 2) frequency-specific suppression, and 3) post-stimulatory suppression in the absence of spiking. We also found a correlative relationship between these nonlinear receptive field features and sensitivity for different rates and directions of FM sweeps, dynamic features found in many vocalizations, including speech. The overwhelming prevalence of cross-frequency interactions revealed by this technique provides further evidence of the central auditory system's role as a pattern-detector, and underscores the need to include nonlinearity in measures of the receptive field.

Keywords

auditory midbrain; mustached bat; frequency modulation; receptive fields; de Bruijn sequence; sideband inhibition; post-excitatory suppression; combination sensitivity; facilitation

Introduction

Frequency modulations (FM) are dynamic information-bearing features found in many animal vocalizations, and are common in consonant-vowel transitions and glides in human

© 2010 Elsevier B.V. All rights reserved.

Corresponding author: W. Owen Brimijoin, MRC Institute of Hear. Res., Queen Elizabeth Building, 16 Alexandra Parade, Glasgow Royal Infirmary, Glasgow UK, G31 2ER, owen@ihr.gla.ac.uk, Phone: 0141 211 4695, Fax: 0141 522 8410.

³Present address: MRC Institute of Hearing Research (Scottish Section), Glasgow Royal Infirmary, Glasgow UK, G31 2ER

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

speech. Discrimination of such sounds depends on correctly identifying the direction and rate of FM in the transitions, and misperception of these features is thought to underlie certain phonological disorders. Within the central auditory system, neurons responsive to FM sounds have been documented in a variety of mammals from the cochlear nucleus to auditory cortex. Such neurons are particularly prevalent in echolocating bats, including the mustached bat (*Pteronotus parnelli*) studied here, presumably because FM sounds are ubiquitous features of both biosonar and social communication sounds. In bats, neurons selective or “specialized” (i.e., exclusively responsive) for FM sweeps occur at all levels downstream of the nuclei of the lateral lemniscus. Moreover, a majority of these neurons are further selective for the direction and rate of frequency shifts in such sounds, analogous to the direction and velocity tuning seen in motion-sensitive visual and somatosensory neurons (Rall, 1964; Segev and Rall, 1998; Warren, *et al.* 1986).

Many studies have implicated interactions between excitation and inhibition as the underlying mechanism for FM sweep selectivity. The classical studies in bats by Suga and colleagues (Suga, 1965; Suga and Schlegel 1973) used a two-tone interaction paradigm to reveal excitatory and inhibitory areas of the receptive field, an approach directly analogous to “two-bar” experiments used to uncover the dynamic interactions between on and off regions of visual receptive fields. They hypothesized that FM directional selectivity depended on the order of activation of these opponent sub-fields.

In a study of FM processing in the inferior colliculus (IC) of the mustached bat, Gordon and O'Neill (1998) showed that directional selectivity for FM sweeps was strongly modulation rate (i.e., sweep velocity) dependent. Using a variation of the two-tone paradigm, they found that the precise temporal and spectral relationship between a best excitatory frequency (BEF) “probe” tone and a best inhibitory frequency “test” tone within the frequency range of the FM sweep could account for rate-dependent directional selectivity (RDS). These findings corroborated similar studies in the pallid bat (Fuzessery 1994). More recently, Brimijoin and O'Neill (2005) extended these findings by systematically varying the delay and frequency of a preceding test tone relative to a fixed BEF probe tone to construct two-tone, non-linear “conditioned response areas.” They then compared the observed spiking responses to FM sweeps of different rates and directions to predictions of the response computed by simple convolution of the FM sweep spectra with the two-tone conditioned-response area, and successfully predicted the actual response to FM sweeps, but only in about 20% of the neurons. Clearly, the two-tone response area was adequate to explain RDS in some neurons, but did not reliably capture the features of the receptive field that are important for determining RDS in the majority of FM responsive IC neurons.

One limitation of this approach that may have accounted for the modest success of these predictions was that the effects of prior test tones were examined only on BEF probe tones. In the present study, we attempted to examine *all* possible two-tone interactions that might occur within an FM sweep. In an effort to improve FM RDS prediction and determine the underlying causes of prediction failure in the two-tone conditioned response area, we introduce the *spectrotemporal response map* (STRM), which assesses spectral and temporal interactions between tone pairs with frequencies arbitrarily spaced in a defined range surrounding the BEF. STRMs are assembled from the responses to a train of contiguous random tones derived from a de Bruijn sequence. The de Bruijn sequence contains a balanced set of all possible tone pair, triplet, and quadruplet transitions, while at the same time retaining a long-term random structure. As such, a de Bruijn stimulus simulates the orderly spectrotemporal structure inherent in FM sweeps, and analyses can be performed on specific tone transitions to explicitly reveal non-linear interactions (e.g., facilitation and suppression) that may underlie specialization for FM sounds. Moreover, by treating the de Bruijn stimulus as a series of random tones and ignoring tone transitions, one can also

construct a “single-tone” STRM, which is mathematically analogous to a *spectrotemporal receptive field* (STRF) constructed from a random-tone sequence using spike-triggered reverse correlation (Valentine and Eggermont 2004; de Charms *et al.* 1998).

Our results show that tone-pair STRMs provide evidence for three types of response properties that are not directly identifiable in linear measures of the receptive field, and that these properties were quantitatively correlated to the observed directional selectivity for some FM sweep rates. Furthermore, these nonlinear response properties may help to explain why conditioned response areas and linear response areas often fail to account for complex RDS for FM.

Methods

Preparation

Two wild-caught adult male Trinidadian mustached bats (*Pteronotus parnelli rubiginosus*) served as subjects in this study. Bats were housed colonially in a temperature-(25° C) and relative humidity-controlled (80%) flight room, and were maintained *ad libitum* on a diet of nutritionally fortified mealworms. Surgery to mount a head-post, implant an indwelling indifferent electrode, and expose the IC was performed under Isoflurane anesthesia (2-3% 2-chloro-2-difluoromethoxy-1,1,1-trifluoro-ethane) using techniques described previously (Brimijoin and O'Neill 2005). Prior to a recording session, bats were injected with a low dose of the short-acting torpogestic Butorphenol (0.7 mg/kg i.p.; Mylan, Canonsburg, PA) and placed in a form-fitting foam body restraint. The head was immobilized in a comfortable position relative to the body by fastening the head-post to a precision, noninvasive stereotactic device specifically designed for bats (Schuller *et al.* 1986). Experiments were conducted on each subject over a 4-6 week period, 4-5 days/week, in 2-4 hour sessions each day.

All husbandry, surgery, and experimental procedures were approved by the University of Rochester's Committee on Animal Resources, and complied with all U.S. federal, state and local ordinances for the humane treatment of research animals.

Recording site and response acquisition

Extracellular recordings were made in the left IC using sharpened glass micropipette electrodes filled with 1M Na⁺ acetate (tip impedances 9-10 MΩ). The microelectrode was lowered into the IC along a dorsoventral trajectory using a piezoelectric micropositioner (Burleigh Inchworm; Burleigh Instrument Division, EXFO, Mississauga, Ont.). The recorded potentials were referenced to a tungsten indifferent electrode surgically embedded just under the skull in contact with the dura mater contralateral to the recording site. The signals were amplified (Dagan, model 2400, Minneapolis, MN), bandpass filtered from 450 Hz to 4.5 kHz using a TDT SC1 filter (Tucker-Davis Technologies, Alachua, FL). Spikes were discriminated based on waveform shape and amplitude with a multi-window TDT SD1 spike discriminator; signal-to-noise ratios were typically 3:1. Spike shapes were continuously monitored to ensure that isolation was maintained. Identified spikes were time stamped (TDT ET1 event timer) and stored along with information about the stimulus in a database (Access©, Microsoft, Redmond WA).

Stimulus calibration and presentation

Sound Stimuli were digitally synthesized with a sample period 3.08 μs in SigGen32© software and Matlab (The Mathworks, Natick, MA), and were presented using TDT System II and III hardware. Sounds were presented through a full-range leaf tweeter (Panasonic model 203) placed 17 cm from the bat's head along the acoustic axis of the pinna at 60 kHz,

25° lateral to midline in the horizontal plane of the stereotactic frame, contralateral to the recording site. The acoustic system's maximum sound pressure level (SPL, dB re 20 μ Pa) was calibrated weekly by placing a 1/4" condenser microphone (Bruel and Kjaer, type 4135) at the position occupied by the bat's head during the experiment. Maximum speaker output ranged from 78 to 83 dB peak equivalent SPL over the 20- to 75-kHz frequency band of the stimuli presented.

Experimental paradigms

Search Stimulus—Most single units in the awake mustached bat IC have little or no spontaneous activity. Therefore, unit isolation was assisted by presenting pure-tone search stimuli generated with a voltage-controlled waveform generator (Wavetek 111, San Diego, CA), gated and shaped into 30-ms tone bursts with a linear rise-fall time of 0.5 ms with a programmable electronic switch (Wilsonics BSIT, San Diego, CA). Sound level was controlled with a programmable attenuator (Wilsonics PATT, San Diego, CA). Once a unit was isolated, its best excitatory frequency (BEF) and minimum threshold were estimated audiovisually.

Frequency Response Areas (FRA)—A coarsely sampled, spectrally broad FRA was initially obtained by presenting tones over a frequency range of 20 to 75 kHz in steps of 1.0 kHz over an amplitude range from 10 dB below the audio-visually estimated minimum threshold up to maximum speaker output in 10 dB steps. The tones were 60 ms in duration with a 0.5 ms rise/fall time, a 120-ms interstimulus onset interval, and were repeated 3-5 times. Based on the tuning width observed in these FRAs, a second, high-resolution FRA was subsequently obtained using tones spaced in 100-Hz intervals over a frequency range spanning the excitatory region of the FRA. Final estimates of BEF and the width of the excitatory area were taken from the high-resolution FRA. The width of the excitatory area was estimated at 20 dB above MT, the level typically used for obtaining responses to stimuli (see below), expressed as a Quality factor, $Q_{20\text{ dB}}$, ($\text{BEF} / \text{bandwidth } 20\text{ dB above MT}$).

FM Stimulus Set—To test the ability of STRMs obtained from tone sequences to predict responses to dynamic, time-varying stimuli behaviorally relevant to the mustached bat, we obtained responses to fixed-bandwidth, upward and downward sweeping FM tones over a range of modulation rates. By necessity, sweep rate and duration co-vary in fixed-bandwidth FM sweeps. Because many IC units can be selective for stimuli of different durations (Brand *et al.* 2000; Casseday *et al.* 1994) it is therefore possible to confound FM sweep rate selectivity with duration selectivity. To control for this, we embedded the sweeps within a four-component "TUDN" stimulus assembly (Brimijoin and O'Neill 2005) consisting of (1) a pure tone at the unit's BEF serving as a control for stimulus duration, (2) an upward sweeping FM, (3) a downward sweeping FM, and (4) a band-limited noise burst. The FM sweeps and the noise bursts had fixed bandwidths of 12 kHz (always ± 6 kHz re BEF), chosen to match the terminal FM sweep in the dominant 2nd harmonic of the mustached bat biosonar signal. As the noise burst had the same long-term spectrum but lacked the orderly temporal structure of the FM sweep, it served as a test of whether responsiveness to the FM sweep might be confounded with a preference for bandlimited noise. Each sound was shaped with a raised cosine envelope (rise/fall time 0.5 ms). The interval between components in the TUDN sequence was 200 ms, yielding a total sequence length of 800 ms. To vary the FM sweep rate, sequences were presented at three stimulus durations: 32, 16, and 8 ms, yielding FM rates of 0.375, 0.750, and 1.50 kHz/ms, respectively. Each stimulus duration/sweep rate condition was presented in separate blocks of 25 repetitions; TUDN sequences were repeated seamlessly within and across blocks.

Pseudorandom de Bruijn Tone Sequence—We synthesized a pseudorandom 9⁴ de Bruijn tone burst sequence (de Bruijn 1946), in which all possible sequential four-tone combinations of nine tone frequencies occur only once. Tone bursts were 4 ms in duration including rise and fall times of 0.5 ms, and were presented sequentially without overlap. Since the tones were shaped with a raised cosine rise-fall time, there was a 1.0 ms gap between tones at peak amplitude and a 0.54 ms gap between tones at 75% amplitude. The resulting sound stimulus consisted of a contiguous train of 6561 tone bursts (total duration = 26.244 s), during which each individual tone frequency (N = 9) occurred 729 times, each possible tone pair (N = 81) occurred 81 times, each possible triplet (N = 729) occurred 9 times, and each possible quadruplet (N = 6561) occurred once. For each unit tested, the de Bruijn sequence was used to generate *de novo* a 24 bit pulse-coded modulation stimulus (wav file) with the center frequency of the stimulus set at the BEF of the unit. In order to match the bandwidth of the FM sweeps and noise bursts presented in the FM stimulus set (see above), the 9 tone frequencies in the sequence were spaced in integer multiples of 1.5 kHz, resulting in a total stimulus bandwidth of 12 kHz. In every case the range of frequencies in the sequence exceeded the bandwidth of the unit's excitatory response area estimated from the FRA (illustrated in Fig. 1A).

The 4 ms tone duration was selected as a compromise that allowed us to reliably drive spiking activity while mimicking the trajectory of the FM sweeps. In principle, an FM sweep can be approximated with a staircase of infinitely short tones sequentially increasing or decreasing in frequency; energy at any particular frequency is only instantaneous (Roverud 1993.) However, the integration time of bat IC neurons sets a tone duration minimum to generate a spiking response that varies from neuron to neuron. The tone duration we chose was short enough to mimic the trajectory of the shortest and longest sweeps using only a pair of tones, but long enough that it elicited robust spiking responses in the IC. Given the 4-ms tone burst duration and the 1.5 kHz frequency interval, the effective modulation rate of the sequential tone transitions ranged from 0.0 to 3 kHz/ms (3 MHz/s), which encompasses the range of FM sweep rates in mustached bat biosonar and communication sounds (Kanwal *et al.* 1994). The de Bruijn tone sequence was presented at the sound level yielding the steepest slope of the BEF rate-level function obtained from the FRA data set; typically this was 10 to 20 dB above minimum threshold. This frequency/amplitude combination in the middle of the unit's dynamic range was chosen because it evoked reliable, but sub-maximal, responses, and was therefore able to reveal both increases and decreases of the BEF response by tone combinations in the sequence. Depending on the responsiveness of the unit, the sequence was repeated seamlessly 4-9 times until a minimum of 1000 spikes were captured.

Data Analysis

FM Stimulus Set—Total spikes were counted in windows extending 64-ms from the onset of each component in the FM stimulus set. For the upward and downward 12-kHz FM sweeps, directional preferences at each modulation rate were quantified with the conventional directional selectivity index (DS) used in many previous studies:

$$DS(U - D)/(U + D) \quad (\text{Eq. 1})$$

where U and D are the spike counts in response to upward and downward FM sweeps, respectively (Brimijoin and O'Neill 2005; Britt and Starr 1976; Gordon and O'Neill 1998; Heil and Irvine 1998; Mendelson and Cynader 1985). Significant directionality was defined as twice the number of spikes in response to one sweep direction over the other, corresponding to an absolute DS index > 0.333.

Single-tone Spectrotemporal Response Maps—Figure 1 illustrates the method for constructing STRMs. First, using BEF and level parameters estimated from the FRA (Fig 1A), tone frequencies were assigned to the de Bruijn sequence, and the stimulus was presented to collect the spike-time data (Fig 1B). Off-line, spike times in a 32-ms window after each occurrence of a given tone in the de Bruijn sequence were collected into separate post-stimulus time histograms (PSTH; 1 ms bin width). The nine resulting PSTHs — one for each of the nine tone frequencies — were stacked from the lowest to the highest tone frequency (Fig. 1C), and the bin counts were converted into a false color map. Each ‘row’ in the final STRM display is the color-mapped PSTH for a different tone frequency as viewed from above: dark blue hues represent low spike counts and bright red hues represent high spike counts (Figure 1D).

Predicting FM Directionality from Single-tone STRMs—First, a frequency-shuffled ‘background’ single-tone STRM was generated using the same spike train, but with a *different* de Bruijn sequence from that was used to generate the tone sequence. This effectively distributes the same spikes over a random set of frequencies while keeping the temporal properties of the STRM intact. This was done to match the background banding pattern, which is caused by the frequent occurrence of excitatory (primarily BEF) tones in the sequence. Second, normalized single-tone STRMs were created by subtracting this background STRM from the uncorrected raw single-tone STRMs. As a consequence, excitation above background resulted in positive values, while suppression below background resulted in negative values. Third, the normalized single-tone STRMs were up-sampled by a factor of 10 in the frequency and time domains to match the frequency/time resolution of the FM sweep spectrograms. Fourth, predictions of the response to FM sweeps were generated by convolving in time the spectrograms of sweeps of both directions at each of three modulation rates from the FM stimulus set with the normalized single-tone STRMs according to standard STRF prediction methods (Andoni *et al.* 2007). Fifth, the predicted response to the sweeps was summed across frequency and half-wave rectified. Finally, DS indices (cf Eq. 1) were calculated for each FM rate using the maximum predicted response for each direction of FM.

Tone-pair Spectrotemporal Response Maps—PSTHs were generated for each of the 81 (9^2) possible tone pairs in the sequence. Spike times were grouped in 1-ms bins in a 32-ms window after each occurrence of the 1st tone in a pair; then the resulting 81 PSTHs were sorted into nine groups by the frequency of that 1st tone. The PSTHs within these 9 groups were stacked low-to-high by the frequency of the 2nd tone in the pair and represented with a false color scale to produce *tone-pair STRMs*. An example tone-pair STRM is shown in Figure 1F.

In principle it is also possible to analyze the responses to tone-triplets and quadruplets. In practice, however, spike counts for these combinations were not high enough for this to be accomplished. For single-tone STRMs, the total number of spikes in response to the sequence is effectively distributed among nine possible tones, whereas for pair analysis, these same spikes are distributed among 81 tone-pairs, further reducing their number. Triplets would essentially divide the number of available spikes by 781 (9^3), rendering meaningful analysis at this and higher levels impractical for all but a few units in the current data set.

Sequence Statistics—The balanced nature of the de Bruijn sequence allows one to isolate single-tone effects within the ongoing response. Because every possible quadruplet of tones is represented exactly once within the sequence, any given tone (or tone-pair) is preceded in each of the three prior frames by an equal number of all possible tones.

Similarly, each tone in the sequence is followed by three frames of equally balanced tones. Consequently, the distribution of tone frequencies for the 12 ms ($3 \cdot 4$ ms) period preceding and following any tone in the sequence is precisely uniform, and any response heterogeneity caused by specific tone interactions within these frames is averaged out. Beyond 3 frames in either direction (a total of 24 ms), however, the relative frequency of any given tone is not perfectly balanced (although their statistics are still roughly uniform), hence the limited time window used for all statistical analyses. All figures depicting STRMs are 32 ms in length; this is purely to allow graphical representation of the response for those units with particularly long response latencies.

Tone-Pair Nomenclature—Given the considerable amount of data produced by this analysis, we adopted a nomenclature to facilitate discussion of specific tone- and tone-pair responses. *BEF* is defined as the frequency that drives the largest spiking response at a given stimulus level, and *tone frequency* will always be specified in *kHz re BEF*. For tone pairs we use the following naming convention [1st tone→2nd tone] to indicate the order of presentation. For example, a transition from a tone 4.5 kHz below the BEF to a tone 3.0 kHz above the BEF is referred to as [-4.5→+3.0]. The top panel in Figure 1F is an assembly of the responses to the tone-pairs whose 1st tone was 6 kHz above the BEF; thus it depicts spiking data for the following pairs (listed from top to bottom): [+6.0→+6.0], [+6.0→+4.5], [+6.0→+3.0], [+6.0→+1.5], [+6.0→BEF], [+6.0→-1.5], [+6.0→-3.0], [+6.0→-4.5], and [+6.0→-6.0]. Similarly, we use shorthand to refer to groups of tone pairs, e.g., [below→BEF] refers to all pairs that begin with a tone below the BEF and end with a BEF tone, while [BEF→x] refers to all pairs whose 1st tone was at BEF.

Predicting FM Directionality from Tone-Pair STRMs—It is not practical to use convolution to predict responses with the higher-dimensional tone-pair STRMs. Instead, we summed normalized PSTHs in response to particular tone pairs having effective modulation rates that matched the FM sweeps in the TUDN stimulus set, i.e., pairs with frequencies lying along the FM sweep trajectories. First, tone-pair STRMs were normalized by the same method used for single-tone STRMs such that excitation above background resulted in positive values, and suppression below background resulted in negative values. Second, relevant pairs were chosen, for example the downward-moving pairs [+3.0→+1.5] and [+1.5→BEF] were used to simulate the 32 ms downward FM sweep. Third, the second PSTH (in this case: [+1.5→BEF]) was shifted forward in time by 4 ms (i.e., the time the actual 32 ms sweep took to change in instantaneous frequency by -1.5 kHz), and the two now temporally-offset PSTHs were summed. In this way, we used four tones to approximate a 32 ms downsweep. Matching pairs were chosen for 32 ms upsweeps and similarly offset and summed. Total spike counts for the 32 ms downward and upward sweeps in a paired set were then used to compute DS indices (via Eq. 1). Pairs used to predict DS indices in response to 16 ms sweeps included [+4.5→+1.5] and [+3.0→BEF] for down FM, and [-4.5→-1.5] and [-3.0→BEF] for up-FM. Similarly the down FM pairs [+6.0→BEF] and [+4.5→-1.5], and up FM pairs [-6.0→BEF] and [-4.5→+1.5] were used to match the modulation rate of the 8 ms sweeps.

Statistical Analysis of Cross-Frequency Interactions—In order to perform a two-way analysis of variance, the spikes in each tone-pair PSTH were summed across a time window beginning at the onset of the tone pair and manually extended for each unit so as to encompass the longest period of facilitation and/or inhibition observed. The total spike counts within these windows were then tallied for each individual repeat of the sequence and analyzed using two-way ANOVA in Prism 4.0 (Graphpad Software, San Diego, CA) for the effects of the tone pairs; i.e., the main effect of the frequency of the 1st tone of the pair, the main effect of the frequency of the 2nd tone of the pair, and their interaction.

Results

We isolated and studied 50 units in the IC of two Trinidadian mustached bats. The mean BEF for these units was 53.9 ± 1.3 kHz (SEM). The lowest BEF recorded was 34.8 kHz and the highest was 59.3 kHz. However, the median BEF was 58.3 kHz, close to the 2nd harmonic of the mustached bat biosonar call. The distribution of BEFs indicates that the vast majority of units were recorded in the prominent dorsoposterior division of the ICC, a large isofrequency region of very sharply tuned neurons with BEFs within 1-2 kHz of the 2nd harmonic, while the rest were recorded in the caudal anterolateral division, which tonotopically represents lower frequencies. Mean excitatory response area bandwidth at MT + 20 dB was 1.5 ± 1.1 kHz (SEM), yielding a mean $Q_{20\text{ dB}}$ of 50.8 ± 1.1 (SEM). Of the 50 units recorded, we collected a complete set of FRAs, TUDN responses, and STRMs for 39 units.

FRAs and single-tone STRMs

We begin by comparing classic FRA representations of auditory receptive fields and single-tone STRMs. The V-shaped FRA in Figure 1A was obtained from a unit with a BEF at 52.4 kHz, which is within the range of the FM component in the 2nd harmonic of the biosonar signal, and was sharply tuned with a $Q_{20\text{ dB}}$ of 33.9 (bandwidth = 1.6 kHz at 28 dB SPL). Alternatively, the BEF from the unit's single-tone STRM (Fig. 1D) was estimated at 52.0 kHz, and the $Q_{20\text{ dB}}$ was 34.7 (bandwidth = 1.5 kHz). These minor differences between the FRA and single-tone STRM estimates of these parameters are attributable to the much coarser resolution of the STRM relative to the FRA (1.5 vs. 0.1 kHz, respectively). For the population, the mean excitatory bandwidth according to the single-tone STRMs was 1.79 ± 0.60 kHz (SEM), vs. 1.55 ± 1.1 kHz for the FRA. Paired t-tests revealed no statistical difference between FRA and STRM estimates for either the BEF ($t_{[49]} = 0.10$, $P = 0.92$) or excitatory bandwidth ($t_{[49]} = 1.08$, $P = 0.28$). While these basic features of the receptive field were closely comparable, it was not unusual for the STRM to reveal other features of the response not captured by the FRA, as discussed below.

Single-tone STRM analysis

The single-tone STRM is constructed by first compiling PSTHs for each of the 9 constituent tones in the de Bruijn stimulus, stacking the PSTHs in order of increasing tone frequency (Fig. 1C), and finally plotting the result in a pseudo-color map, as if viewing the histograms from above (Fig. 1D). A notable temporal feature evident in the "raw" STRM is a low-level banding pattern repeating at 4 ms intervals, which is the duration of each tone, and therefore the inter-tone interval or "frame time", in the presented sequence. The banding is generated by the relatively frequent occurrence of suprathreshold excitatory tones in the de Bruijn sequence, the most significant of which are BEF tones that on average occur 1/9th of the time. The resulting rhythmic driven activity fortuitously provides a backdrop for revealing suppression that would otherwise be invisible, because spontaneous activity is typically very low in mustached bat IC neurons. In fact, the unit in Figure 1 had no detectable spontaneous activity (see the FRA in Fig. 1A).

The single-tone STRM for the unit in Figure 1D revealed a phasic excitatory BEF response (bright spot) followed in time by suppression (dark region), and this was the typical pattern in the units we recorded. Approximately 70% of the units also had inhibitory sidebands, revealed by a gap in background banding following the BEF response at higher and/or lower frequencies. Note that the dark regions temporally concurrent with the BEF response are not necessarily indicative of sideband suppression; the sequence has no simultaneous tones, making it unable to differentiate simultaneous masking from simple lack of excitatory drive. The unit shown in Figure 1D, however, did show sideband suppression below and following

the BEF response. Inhibitory sidebands, as well as other sources of spectrotemporal asymmetry to be discussed later, have been implicated by many previous studies as a possible cause of FM directional selectivity (Brimijoin and O'Neill 2005; Fuzessery and Hall 1996; Fuzessery *et al.* 2006; Gordon and O'Neill 1998; Suga and Schlegel 1973).

Tone-Pair STRMs Reveal the Form and Prevalence of Cross-Frequency Interactions

The de Bruijn sequence was chosen because it includes ordered tone patterns, the simplest of which are tone pairs. Two-tone stimuli have traditionally been employed to study the mechanisms underlying responses to FM, because the discrete spectral transitions in tone pairs approximate the spectrotemporal transitions in the FM sweep. Two-way ANOVA of pair-wise spike counts for all units showed that responses to each tone in a pair were not independent. The main effect of 1st tone frequency was significant ($F_{(1,8)} = 198.6, P < 0.05$), the main effect of 2nd tone frequency was significant ($F_{(1,8)} = 207.2, P < 0.05$), and the interaction was significant ($F_{(1,64)} = 11.3, P < 0.05$). The main effects accounted for 70.0% of the variance, whereas interactions contributed less and were highly variable, accounting for 19.0% of the variance.

We identified five non-exclusive classes of two-tone interactions: 1) *sideband suppression*, 2) *post-excitatory suppression*, 3) *two-tone facilitation*, 4) *post-stimulatory suppression*, and 5) *frequency-specific suppression*. Sideband suppression, defined as the ability of off-BEF tones to suppress responses to BEF tones, was found in 70% of the units in our sample. Post-excitatory suppression (PES), typically defined as a period of suppression after the excitatory response to BEF tones (but see below), was also common, occurring in 82% of the units. Most instances of sideband- and post-excitatory suppression were also visible in single-tone STRMs as below-background activity. Since one of the primary goals of this experiment was to reveal cross-frequency interactions not evident in such linear measures, we did not examine these basic interactions further. The nature, prevalence, and predictive power of the three remaining classes of interactions are examined in turn, and each is accompanied by an example unit.

Two-Tone Facilitation—Twenty-two percent of the isolated units had excitatory regions in their single-tone STRMs that were tilted in the spectrotemporal domain. However, it is not possible to determine from single-tone STRMs whether such a pattern is simply due to excitatory responses at different latencies for different tone frequencies, or instead reflects non-linear interactions between tones (combination sensitivity). By contrast, tone-pair STRMs differentiate between these two possibilities, and they showed that units with tilts attributable to simple frequency-dependent latency shifts were in the minority, constituting 25% of the cases. The tilt in the remaining 75% resulted from time-dependent two-tone facilitation.

The unit shown in Figure 1 provides a good example of super-additive two-tone facilitation. The unit's single-tone STRM (Fig. 1D) shows that the BEF spiking response occurred with a latency of 7 ms. This excitation was flanked on the low-frequency side by sideband suppression, and was followed by post-excitatory suppression. The single-tone STRM also showed a very weak excitatory response at +3 kHz with a latency of 11 ms, a feature not visible in the FRA (Fig. 1A). Whether the responses to the BEF tone and this higher tone were independent or not is impossible to determine from the single-tone STRM.

However, the tone-pair STRMs (Fig. 1F) are able to discriminate between these two possibilities. If the two responses were independent, then the tone-pair STRMs should show a spiking response to all pairs beginning with either BEF or +3 kHz tones. However, this was not the case. Figure 1F shows all of the 81 tone-pair responses grouped by 1st tone frequency. Beginning with the top panel, the “hot spot” at the center of the display indicates

a strong response to [+6.0→BEF] pairs with a latency of 11 ms. This latency is equal to the single-tone BEF response latency of 7 ms plus one 4-ms tone-frame duration, suggesting that the response is simply due to the 2nd tone in the pair, namely the BEF tone. No other tone following a +6.0 kHz tone evoked an above-background response, including +3 kHz tones, which were very weakly excitatory in the single-tone STRM. We would conclude from this panel that +6 kHz tones had little measurable effect on the responses to any of the other tones in the sequence.

Other tone pair groupings showed a pattern similar to that found in the top panel. The second and last (bottom) panels of Figure 1F show that the response to pairs that start respectively with either +4.5 kHz or -6.0 kHz tones was essentially identical to the response to pairs that start with +6.0 kHz tones. The similarity of these tone-pair STRMs to the single-tone STRM (Fig. 1D) suggests that these three off-BEF tones have no effect on responses to subsequent tones in the range tested.

On the other hand, 1st tones at BEF significantly affected responses to all 2nd tones, as shown in the center panel in Figure 1F. This panel shows a vertical band of excitation with a latency of 7 ms, and the homogeneity of this band demonstrates that 1st tones at BEF were excitatory regardless of which tones followed them (i.e., they were not backward-masked). Note that leading BEF tones suppress the response to trailing BEF tones, which would be seen as excitation at BEF with a latency of 11 ms, similar to those corresponding regions in the top and bottom panels of the figure. This result is consistent with the on-BEF suppression seen in the single-tone STRM (Fig. 1D).

Sideband suppression can be seen in the sixth and seventh panels of Figure 1F. All 2nd tone responses were suppressed when preceded by 1st tones at either -1.5 kHz or -3.0 kHz. The growing activity at BEF in the lower two panels shows the progressive weakening of low-side suppression with increasing separation from BEF.

While the interactions described above are not much different from those visible in the single-tone STRM, the tone-pair STRM depicted in the third panel provides evidence that the apparent response to +3 kHz tones is indeed attributable to two-tone facilitation, not frequency-dependent shifts in latency. Recall that the single-tone STRM (Fig. 1D) showed that +3 kHz tones were weakly excitatory with a latency of ~11 ms, even though +3 kHz tones fell outside the excitatory response area in the FRA. In the tone-pair STRM, if +3 kHz tones were excitatory on their own, we would expect a vertical band of excitation in the third panel, similar to that seen for BEF in the center panel, but at a latency of ~11 ms. However, no such band is seen, which proves that +3 kHz tones in fact generated no excitatory response on their own. Instead, a powerful facilitated BEF response totaling 341 spikes is seen to [+3 kHz→BEF] *pairs*, far exceeding the sum of the responses to BEF and +3.0 kHz 1st tones alone (175 spikes). This example shows how tone-pair STRMs are able to distinguish between super-additive (non-linear) and additive (linear) responses, whereas single-tone STRMs cannot.

While this example illustrates the analytical advantage of the tone-pair STRM for revealing non-linear interactions, the question remains whether this advantage carries over to the prediction of directional selectivity to FM sweeps. Figure 1E shows the summed responses to tones, FM sweeps, and narrowband noise from the TUDN stimulus. This unit responded well to BEF tones at all tested durations (*dashed line, open circles*). By contrast, there was no response to narrowband noise (*dashed line, open squares*) or upward FM sweeps (*solid line, filled upward triangles*) of any duration. However, whereas the unit did not respond to 32 ms-duration (i.e., slow) downward FM sweeps (*solid line, filled downward triangles*), it was strongly responsive to more rapid downward FM sweeps of 16 and 8 ms duration. The

response to such fast downward FM sweeps was clearly facilitated compared to the response to tones of the same duration. Based on delay-tuned sideband-inhibitory models of FM sweep selectivity (Brimijoin and O'Neill 2005; Gordon and O'Neill 1998; Suga and Schlegel 1973; Suga *et al.* 1974), the below-BEF suppression dominating the STRMs would account for the lack of response to either narrowband noise or upward FM sweeps. Moreover, the modulation rate of the preferred 16 ms downward FM sweep (-0.75 kHz/ms) matches the instantaneous rate of frequency change found in the highly facilitative $[+3.0 \rightarrow \text{BEF}]$ tone pairs in the two-tone STRM. This two-tone non-linear spectrotemporal interaction could account for facilitation by rapid downward FM sweeps.

Post-Stimulatory Suppression—In the vast majority of units in our sample, 1st tones at BEF tended to reduce or eliminate the spiking response to 2nd tones at BEF. This phenomenon resembles post-excitatory suppression (PES) in auditory nerve fibers, where it is classically attributed to adaptation at the hair cell/primary fiber synapse (Westerman and Smith 1984). At the level of the ICC, other factors such as synaptic inhibition could also be contributing to this effect. Fully 82% of units in our sample displayed some degree of on-BEF suppression, while approximately 45% showed no response to 2nd tones at BEF when immediately preceded by BEF tones. STRM analysis revealed that all of these fully-suppressed units also showed suppression following 2nd tones at BEF even when these tones did not elicit a supra-threshold (spiking) response. Because this effect clearly differs from classically-defined PES (which involves adaptation), we term this *post-stimulatory suppression* (PSS).

The unit represented in Figure 2 illustrates the features of PSS, and again shows the power of the tone-pair STRM in revealing possible mechanisms. The FRA (Fig. 2A) shows a classic v-shape tuning, with a BEF of 58.2 kHz. The single-tone STRM (Fig. 2B) shows that the unit responded strongly to BEF tones with a latency of ~ 7 ms. Tones 1.5 kHz above and below the BEF were also excitatory, but much less so. Each of these three frequencies was also followed by a period of PSS, the duration of which appears to be roughly correlated with the activity preceding it.

We might conclude from this that the spiking response is directly responsible for PSS. But the tone-pair STRMs (Fig. 2D) reveal that PSS was not dependent on spiking. As in Figure 1F, the homogeneous vertical band in the center panel indicates that all tone pairs starting with BEF tones were excitatory with a latency of ~ 7 ms. This response was followed by complete suppression at 11 ms, showing that BEF 1st tones fully suppressed responses to all subsequent tones, including BEF. However, the prolonged suppression in the center of the vertical band at 15 ms (red arrow) argues that 2nd tones at +1.5 kHz, BEF, and -1.5 kHz extended the period of suppression even though they were themselves suppressed by 1st BEF tones and showed no spiking response. That there is no suppression of the background response at 15 ms associated with other $[\text{BEF} \rightarrow x]$ pairs argues that the extended PSS cannot be attributed to persistent suppression by BEF 1st tones. Given the placement and size of the suppressive notch, the data suggest that PSS is elicited by otherwise excitatory 2nd tones, regardless of whether they actually generate spiking activity. Note that this feature is also visible for $[\text{BEF} \rightarrow \text{above}]$ pairs for the unit in Figure 3.

As in the previous example, the other features of this unit's tone-pair STRMs are in qualitative agreement with responses to FM sweeps. Like the previous example, this unit was directionally selective for fast downward FM sweeps (Fig. 2C). The modulation rate of the preferred 8 ms sweep (-1.5 kHz/ms) matches the effective modulation rate of $[+6.0 \rightarrow \text{BEF}]$ tone pair transitions, which elicited the greatest response in the tone-pair STRMs (Fig. 2D, top panel). Similarly, the modulation rate for 16 ms sweeps (-0.75 kHz/ms), which elicited just slightly less response than 8 ms sweeps, matches the effective rate

for [+3.0→BEF] pairs (Fig. 2D, 3rd panel), which were nearly as potent as [+6.0→BEF] pairs. The response to upward FM sweeps is much lower overall, and this is reflected in the lower spike counts for [below→BEF] tone pair transitions in the tone-pair STRM. Correspondingly, the worst upward FM sweep modulation rate (+0.375 kHz/ms) matches the effective modulation rate of [-1.5→BEF] pairs, the [below→BEF] pair with the lowest spike count.

By contrast to the unit shown in Figure 1, which clearly was facilitated by downward sweeps above all other stimuli including BEF tones, the responses to the preferred downward FM sweeps (*downward triangles*) in this unit were about the same as that to BEF tones (Fig. 2C, *open circles*). The preference for downward sweeps comes about here not by facilitation in the preferred sweep direction, but by suppression in the non-preferred direction, which was by far the more common mechanism for sweep direction preference seen in a previous study of the mustached bat IC (Gordon and O'Neill 1998). It should also be noted that the apparent preference for faster downward FM sweeps in the unit shown in Figure 2 may actually reflect a preference for brief stimuli, as there is clearly an inverse correlation between duration and response for all stimulus types (Fig. 2C). In particular, the curves for BEF tones and downward FM sweeps were closely matched. One explanation for this preference for shorter stimuli may be the prolonged PSS to excitatory tones (e.g., at BEF) shown in the tone-pair STRMs. This suppression persists twice as long as the 4-ms tone duration used in the de Bruijn stimulus. If the period of suppression grows proportionately with stimulus duration, it might result in a buildup of forward masking as the duration (and therefore, duty cycle) of the components in the TUDN stimulus increases.

Frequency-Specific Suppression—Suppression, either in the form of sideband inhibition or PSS, was found in 94% of the units. In at least 20% of these units, suppression was frequency specific, affecting the response to some frequencies more strongly than others. Most commonly, the suppression following a 1st tone tended to be stronger for 2nd tones of a similar frequency. Given that such frequency-specific suppression can only be seen when the response area is wider than the 1.5 kHz spectral interval of our sequence, and that only 23% of the units had response areas wider than this, it is impossible to estimate the true prevalence of this type of suppression.

One such unit is shown in Figure 3, recorded in the ventral IC. The unit was broadly tuned with a BEF of 47 kHz (Fig 3A). The single-tone STRM (Fig. 3B) shows an excitatory response region skewed by a high-frequency tail extending from the BEF to +4.5 kHz and a fast response latency of ~5 ms. Strong, brief suppression follows BEF tones, while progressively weaker suppression follows the less-excitatory tones above the BEF. This correlation between the strength of excitation and the strength and duration of suppression appears similar to that observed in the previous unit in Figure 2B, but recall that excitation per se is not necessarily causally related to such suppression, which may arise from inhibitory activity from a separate input.

The tone-pair STRMs in Figure 3D clearly show this gradient of early excitatory responses to 1st tones from BEF up to +4.5 kHz, mirroring the excitation pattern seen in the single tone-STRM. Following below-BEF 1st tones, the high-frequency tail is also visible for 2nd tone responses over this same frequency range. On the other hand, following above-BEF 1st tones, the high-frequency tail is limited to frequencies immediately above BEF. Comparing the response to [-1.5→+1.5] pairs with [+1.5→+1.5] pairs, it is clear that the response to +1.5 kHz 2nd tones is more strongly suppressed by +1.5 kHz 1st tones than by -1.5 kHz 1st tones (*red arrows*). A similar asymmetry may be observed for [-3.0→+3.0] and [+3.0→+3.0] pairs (*white arrows*). These data suggest that the high-frequency tail of this neuron's receptive field was more strongly suppressed by above-BEF 1st tones, i.e., those most

similar in frequency to the high-frequency tail. Note that this type of frequency-specific suppression is also visible in the example unit in Figure 2D in the selective suppression of -1.5 kHz 2nd tones by -1.5 kHz 1st tones - and to a lesser extent, the suppression of +1.5 kHz 2nd tones by +3.0 kHz 1st tones.

The asymmetry in this unit's single-tone STRM predicts a good response to very rapid downward FM sweeps. The actual FM response shown in Figure 3C indicates that the unit was in fact directionally selective for upward FM by virtue of suppression by downward FM, but had little to no rate preference. This finding is in qualitative agreement with the tone-pair STRM results, as we found greater suppression with [above→BEF] pairs than [below→BEF] pairs. The greatest response in the tone-pair STRMs was found for [-6.0 kHz→BEF] tone pairs. The effective modulation rate for such pairs (+1.5 kHz/ms) corresponds to an 8 ms upward FM, which was (marginally) the best duration and sweep direction in the FM stimulus set.

STRM-predictions of response to frequency modulations

Single-tone STRM-predictions—To test the ability of the single-tone STRM to predict responses to time-varying signals such as FM sweeps, we convolved the single-tone STRM with the spectrograms of the FM stimuli from the TUDN stimulus set to generate predictions of FM responses. Directional selectivity indices (*DS*) for the three FM sweep rates tested were computed from these predictions and cross-correlated with the observed FM directionality for the same sweeps in each unit, as shown in Figure 4A. Points lying along the *diagonal line* with a slope of 1.0 indicate that the single-tone STRM accurately predicted FM directional selectivity in these units. However, whereas the observed *DS*'s ranged widely from -1.0 to +1.0, those predicted from the single-tone STRM fell mostly in a much more compressed range from -0.33 to +0.33, greatly underestimating the observed magnitude of directionality. In fact, single-tone STRMs predicted significant directionality (*DS* ≥ 0.3) in only 4 sweep rate/unit combinations, whereas 32 rate/unit combinations were found to be significantly directional using FM sweeps. Thus the single-tone STRM is unreliable as a predictor of FM directionality.

Tone-pair STRM-predictions—No attempt was made to adapt convolution prediction methods such as those used to generate Figure 4A to the tone-pair STRMs, as the higher dimensionality of the tone-pair STRMs renders such analysis intractable. Instead, simplified predictions of FM directionality were generated in the form of tone-pair STRM asymmetry. Figure 4B illustrates the correlation between *DS* indices generated from tone-pair STRMs and those measured from actual FM responses. The filled squares represent *DS* indices for slow (32 ms) FM sweeps, the open circles 16 ms sweeps, and the asterisks 8 ms sweeps. Overall, the tone-pair STRMs predicted greatly more directionality than did the single-tone STRMs. For 32 ms sweeps, predicted *DS* indices were correlated with actual *DS* indices with an R^2 value of 0.47, for 16 ms sweeps the R^2 dropped to 0.01, and the correlation for 8 ms sweeps was equal to 0.22.

DISCUSSION

Using a train of tones whose order was specified by a balanced de Bruijn sequence proved to be a powerful technique for uncovering non-linear interactions potentially underlying responses to structured, spectrotemporally dynamic sounds. In some cases, the nonlinear interactions help to explain why using conditioned response areas and/or linear receptive fields to predict FM responses can fail. It allowed us to quantify the effect of multi-tone interactions efficiently, and revealed a degree of complexity in the receptive field that cannot be gleaned from conditioned response areas or linear measures of the receptive field.

The tonal interactions we observed were qualitatively in agreement with the differential responses to FM sweeps in the example units. A significant correlation was observed between predicted versus actual DS indices for the slowest FM rate. Three general types of multi-tone interactions were documented in the current study; each of these particular interactions and their implications are discussed below.

Two-Tone Facilitation

Two-tone facilitation is a commonly observed feature of the mammalian central auditory system (Brimijoin and O'Neill 2005; Brosch and Schreiner 2000; 1997; Fitzpatrick *et al.* 1993; Peterson *et al.* 2008; Leroy and Wenstrup 2000; Mittmann and Wenstrup 1995; Olsen and Suga 1991; Portfors and Wenstrup 2002; Sanchez *et al.* 2008; Suga *et al.* 1978). The conditioned response area is entirely capable of quantifying two-tone facilitation (Brimijoin and O'Neill 2005), so in this instance, the current study did not yield insight into the FM-prediction failures of previous methods. The prevalence of this nonlinear property, however, has implications for convolution-based methods of response prediction. A number of studies have noted frequency-dependent latency shifts in excitatory areas of linear STRFs, as we found in 22% of the single-tone STRMs in our unit sample. This inseparability, sometimes described as spectrotemporal tilt when found in contiguous regions of the receptive field, has been hypothesized to account for FM rate and/or direction selectivity (Atencio *et al.* 2007; deCharms *et al.* 1998; Depireux *et al.* 2001; Linden *et al.* 2003; Miller *et al.* 2002; Sen *et al.* 2001). Our results show that in the mustached bat IC, spectrotemporal tilt is more often the result of facilitation of the BEF response (75% of cases) rather than simply template matching along the spectrotemporal trajectory of a frequency-dependent latency shift (25% of cases).

An inability to discriminate between additive (linear) and super-additive (non-linear) interactions that result in a given single-tone STRM's spectrotemporal tilt could give rise to erroneous predictions. If an excitatory region in a single-tone STRM is the result of linear processes, then adding energy to that region should result in an increase in spike rate. However, if the excitatory region is the result of frequency *a* facilitating in a time-dependent manner the response to frequency *b*, then tones presented at frequency *a* in isolation from *b* may not increase spike rate. Furthermore, the two essential frequencies *a* and *b* will not increase spike rate unless they are presented in the preferred temporal relationship. In short, a particular spectrotemporal *pattern* of stimulation is required to drive the facilitated response. As a linear measure, single-tone STRMs cannot differentiate between facilitation and excitation, let alone quantify the degree of facilitation, simply because the majority of spiking activity is the result of energy in the excitatory portion of the receptive field - and energy at the BEF occurs relatively rarely in the proper temporal relationship with energy at the facilitating frequency. That is, for random tone stimuli, the specific tone combinations generating non-linear interactions occur only a small (and undocumented) fraction of the time relative to each tone alone or when the two tones occur in an unfavorable temporal relationship. Indeed, facilitation cannot be accurately quantified in *any* stimulus whose specific transitional probabilities are not explicitly identified. Because so many neurons at midbrain and higher levels exhibit such non-linear spectrotemporal interactions, it is perhaps not surprising that linear techniques often fail to predict responses to biologically relevant complex sounds containing these patterns (Machens *et al.* 2004).

Facilitation involving widely spaced frequencies separated by one or more octaves can be disrupted by selectively blocking glycinergic inputs to distal dendrites of combination sensitive neurons in mustached bat IC (Peterson *et al.* 2008; Sanchez *et al.* 2008). The facilitation described here and in a previous study (Brimijoin and O'Neill 2005) involves tones separated by much smaller spectral intervals, and the synaptic mechanism is unknown.

However, the long latency of the example unit in Figure 1 (and other units showing facilitative interactions) is consistent with the hypothesized distal dendritic locus of facilitation, as the facilitated synaptic potentials would reach the soma slightly later than those from more proximal excitatory inputs.

Post-Stimulatory Suppression

Nearly all studies of the IC have found that responses to BEF tones exhibit both excitatory and inhibitory components, though it is usually unclear whether these arise from single or multiple (possibly independent) afferent inputs, local circuits, or intrinsic membrane currents. Results from the present study suggest that in many cases these components may arise independently from two or more sources. Specifically, we showed that BEF 1st tones can suppress the excitatory response to 2nd tones at and near the BEF, while leaving intact the suppression caused by the 2nd tone. This nonlinear response property was seen in *all* units with on-BEF suppression. The previously employed *conditioned response area* (Brimijoin and O'Neill, 2005) cannot detect this nonlinear response property, and this may account for some FM prediction failures in that model. For slow FM sweeps, which contain energy at and near the BEF over a longer period, such suppression may be important for shaping a neuron's response as it may work over a longer time course than that predicted by either two-tone masking or linear receptive fields.

A similar phenomenon has been observed psychophysically: tones whose audibility is reduced by a preceding masker may still themselves reduce audibility of a third tone (Plack *et al.* 2006). Figure 5A illustrates a basic model that explains this finding by postulating that the masking is due to a delayed (or concurrent but prolonged) inhibitory input whose tuning overlaps that of a separate excitatory input; this is a concept not without precedence (Sutter *et al.* 1999; Wehr and Zador 2003). In this scenario, the delayed and/or prolonged suppression caused by a leading BEF tone would collide with the excitation caused by a trailing BEF tone, reducing or eliminating the excitatory response. Delayed/prolonged inhibitory input is the linchpin of inhibitory sideband models for FM rate and directional preference (Fig. 5A, *left*) (Fuzessery 1994; Gordon and O'Neill 1998; Heil *et al.* 1992; Shannon-Hartman *et al.* 1992; Suga and Schlegel 1973; Suga *et al.* 1974), albeit in this case with overlapping ("on-BEF") excitatory and inhibitory tuning (Fig. 5A, *right*). Accordingly, post-stimulatory suppression is the result of the delayed/prolonged suppression caused by the 2nd BEF tone.

The short temporal gap between contiguous tones was smaller than the minimum detectable gap typically found in the mouse IC (Walton *et al.* 1997); this raises the question of whether PSS is simply attributable to the unit treating 3 consecutive tones (e.g.: BEF→BEF→BEF) as a single 12 ms tone and responding only at the onset. PSS, however, was also seen following [BEF→+1.5] and [BEF→-1.5] pairs (see Figure 2). We feel that given the evidence for PSS following pairs composed of different frequency tones, the delayed inhibition in the model in Figure 5A remains the most parsimonious explanation for this phenomenon, as it accounts for the lack of response seen following both same-frequency and different-frequency tone pairs. Thus sideband (lateral) inhibition, PSS, and on-BEF suppression appear to be mechanistically related, and can be modeled as manifestations of the same fundamental pattern of neuronal circuitry.

It is in suggestions of anatomical connectivity that the STRM method proves useful in guiding the formation of hypotheses. Inhibition in the IC arises from intrinsic circuitry, ascending brainstem projections, and possibly corticofugal inputs, which exert their effects by turning off excitatory inputs to the IC (Covey *et al.* 1996; Nataraj and Wenstrup 2006; Peterson *et al.* 2008). The data here suggest that the responses of neurons demonstrating post-stimulatory suppression are dependent on a mixture of both types of inhibition (local

and distant). These IC neurons likely receive an independent delayed inhibitory input from one source, as well as an excitatory input from another source that itself can be sideband inhibited. Sideband inhibition is common in the ventral and dorsal divisions of the cochlear nucleus (Rhode and Greenberg 1994), as well as the intermediate nucleus of the lateral lemniscus (Portfors and Wenstrup 2001), all of which strongly project to the IC.

Frequency-Specific Suppression

Suppression in a linear receptive field is treated by convolution-based prediction methods as a non-selective negative gain masking responses to all subsequent acoustic features. Similarly, simple forward masking of BEF responses, such as that measured using two-tone CRAs, cannot detect the presence of frequency-specific suppression. Our tone-pair STRM results suggest that suppression, whether sideband or post-stimulatory, can in many cases selectively target particular frequencies. As a result, a suppressive region in a two-tone CRA or a linear receptive field may be over-weighted in convolution-based predictions, because that region is not suppressive when subsequent acoustic features are sufficiently different in frequency. This overestimation may be particularly applicable to predictions of FM responses, since instantaneous frequency is constantly changing in a rate-dependent manner, traversing regions of the receptive field over time.

Unlike the model for post-stimulatory suppression, an anatomical connectivity model giving rise to frequency-specific suppression requires more than two components. The most parsimonious would be a two-stage model consisting of two inputs to the cell (labeled 1 and 2 in Fig. 5B) each with its own set of excitatory and delayed-inhibitory inputs (like those in the model for post-stimulatory suppression). The model predicts that, depending on the frequency of the masking tone and the tuning of the two inhibitory arms of the model, one or both of the excitatory inputs could be silenced. Such a condition would affect the shape of the receptive field for 2nd tones in a frequency-dependent manner. A classical role for such frequency-specific suppression may be echo suppression (Wickesberg and Oertel 1990), since echoes consist of a delayed repeat of a spectrally similar sound. The model is in agreement with results from intracellular recordings in the IC showing that inhibition is not always local, but may be “inherited” from brainstem inputs that inactivate separate excitatory inputs (Nataraj and Wenstrup 2006; Peterson *et al.* 2008).

STRM Predictions

Single-tone STRMs may contain features such as asymmetrical inhibitory sidebands and spectrotemporal tilt that might be useful for predicting responses to FM sweeps. FM sweep responses have proven difficult to predict from linear receptive fields (Barbour and Wang 2003; Machens *et al.* 2004) (and see Fig. 4A), though success has been achieved predicting slow broadband FM responses in the owl monkey auditory cortex (Atencio *et al.* 2007) and directionality in response to drifting spectral ripples in the freetailed bat IC (Andoni *et al.* 2007). Non-linear two-tone CRAs proved somewhat better in predicting responses to FM in the IC (Brimijoin and O'Neill 2005), although successful predictions for both sweep directions and all tested sweep rates were achieved in only about 20% of the units. The predictive power of all these techniques is complicated by such factors as species differences, anatomical location, and the overall prevalence of non-linear response properties like those described above. The tone-pair STRMs were somewhat more successful in predicting FM directionality than were single-tone STRMs. Tone-pair STRMs in general predicted greatly more directionality than did the single-tone STRMs and directionality predictions were correlated with actual directionality at least for the slowest FM sweeps. There was a more modest correlation between predicted and actual DS indices for the fastest 8 ms sweeps, and no correlation was found for 16 ms sweeps.

STRM Prediction Failure

The primary cause of the single-tone STRM's prediction failure can be traced to its underlying assumption of linearity, i.e., that the spiking response to sound at one frequency is largely or entirely independent of the response to sounds at other frequencies and times. It is recognized that the response of neurons in the central auditory system is dependent on stimulus context (Ahrens *et al.* 2008; Christianson *et al.* 2008). For example, two-tone interaction experiments have found cross-frequency, and often strongly time-dependent suppressive, sub-additive, or supra-additive (facilitatory) interactions in the responses of central auditory neurons in a wide variety of species (Calford and Semple 1995; Faure *et al.* 2003; Gordon and O'Neill 1998; Harris and Dallos 1979; Kaltenbach *et al.* 1992; O'Neill and Suga 1979; Suga and O'Neill 1979; Suga *et al.* 1983; Brimijoin and O'Neill 2005; Brosch and Schreiner 2000; 1997; Mittmann and Wenstrup 1995; Portfors and Wenstrup 1999). Linear receptive fields have often been shown to lack generalization beyond the stimulus used to generate them (Klein *et al.* 2000; Sen *et al.* 2001; Theunissen *et al.* 2000; Valentine and Eggermont 2004; Woolley *et al.* 2006). This limitation impacts their utility for predicting responses to novel sounds (Bar-Yosef *et al.* 2002; Kowalski *et al.* 1996).

FM predictions generated from *tone-pair* STRMs were more accurate than those generated from single-tone STRMs, but in general the tone-pair STRM does not greatly surpass the conditioned response area in prediction accuracy (Brimijoin and O'Neill, 2005). Given that a large number of errors arose for faster FM sweeps, we believe that the effective temporal sampling period of 4 ms and frequency spacing of 1.5 kHz was too coarse. Evidence for an error related to temporal sampling can be found in the unit in Figure 1. This unit responds well to sweep rates faster than those matching the “best tone pair” [+3.0→BEF]. The reason for this discrepancy may lie in the fact that these most-rapid sweeps pass through the +3.0 kHz facilitatory region 2.0 ms prior to sweeping through the BEF region, a time interval that may also be facilitatory. Facilitation by +3.0 kHz tones thus appears to have an effective time window extending from (at minimum) 2 to 4 ms. Our (longer) temporal sampling period of 4 ms precludes temporal measurements at a finer temporal scale than this. Spectral sampling may also be a problem because most of the neurons were too narrowly tuned to allow more than one tone within the excitatory portion of the receptive field. This accounts for the inability to precisely determine the overall prevalence of frequency-specific suppression. Other errors in prediction may be due to the fact that this temporal and spectral sampling constraint also applies to measures of inhibition, or to a fundamental limitation encountered by attempting to approximate a frequency sweep with a set of discrete tones.

A final, but related qualification of the STRM methodology is that two-tone STRMs are not ideal predictors of response to complex sounds other than FM. It is not possible, for example, to use them to predict responses to communication sounds or even narrowband noise since the de Bruijn tonal sequence never contains simultaneous energy at more than one frequency. To measure simultaneous cross-frequency interactions in addition to forward/backward masking effects with a de Bruijn sequence would have statistical requirements would render it too long in duration to be experimentally feasible. Such limitations also render it impossible to quantify simultaneous facilitation, a nonlinear response property found in the majority of units in the ICC (Mittman and Wenstrup 1995; Portfors and Wenstrup 2001), as well as large populations of neurons in the medial geniculate body (Olsen and Suga 1991) and auditory cortex (reviewed by O'Neill 1995). Tone-pair STRMs are also not a good predictor of pure tone duration-sensitivity, as they are measured using discrete tones that can invoke individual onset responses.

Pattern-Detection and the Receptive Field

An important advance in our understanding of the receptive field came from the use of spike-triggered reverse correlation techniques to define the STRF (Aertsen *et al.* 1980; Eggermont *et al.* 1981). The STRF samples the receptive field in the dimensions of frequency and time to capture how the filter properties of auditory neurons change as a function of time. Auditory STRFs have been measured using tone or multi-tone chord sequences, as well as a variety of static or dynamic broadband ripple stimuli similar to the drifting contrast gratings used in studies of visual receptive fields (deCharms *et al.* 1998; Escabi and Schreiner 2002; Klein *et al.* 2000; Kowalski *et al.* 1996; Rutkowski *et al.* 2002). They have also been estimated with natural sound stimuli such as vocalizations and environmental sounds (Machens *et al.* 2004; Sen *et al.* 2001; Theunissen *et al.* 2000). STRFs have considerably advanced our knowledge of spectral and temporal integration in the central auditory system, providing insight into which higher-order features of ethologically-relevant sounds are captured (deCharms *et al.* 1998; Theunissen *et al.* 2004) and how the receptive field is modulated by behavioral contingency, attention, and context (Fritz *et al.* 2003; Fritz *et al.* 2005; Kvale and Schreiner 2004; Shechter and Depireux 2007). Augmenting this powerful technique with measures of spectrotemporal pattern-sensitivity could potentially improve response predictions and enhance our understanding of neural responses to natural sounds.

The tone-pair STRM derived from the de Bruijn sequence revealed features of the response to tone pairs that cannot be ascertained directly from either conditioned response areas or linear single-tone STRMs. This novel technique succeeds because many (perhaps most) IC neurons are non-linear pattern detectors. It has been argued that central auditory neurons are adapted by experience to process the time/frequency statistics of the particular sounds to which they are exposed (Singh and Theunissen 2003), a process that may even be manifested in primary auditory nerve fibers (Rieke *et al.* 1995). Natural sounds contain only a small subset of the higher-order statistics found in truly random stimuli, and it is not surprising that linear receptive fields estimated using natural sounds have often been found to be quite different from those derived from random tones (Machens *et al.* 2004; Sen *et al.* 2001; Theunissen *et al.* 2000). Natural stimuli are more likely to engage time-dependent cross-frequency non-linear interactions because they contain a far higher percentage of the particular spectrotemporal patterns that have shaped response properties through experience and evolution. Using tonal sequences that contain identifiable frequency transitions - including those found in naturalistic spectrotemporal patterns - could be invaluable in increasing our understanding of the encoding of complex sound in the central auditory system.

Acknowledgments

This research was supported by a grant from the National Institute for Deafness and Communicative Disorders (NIDCD R01-DC03717) and from a National Institute of Mental Health Training Grant (5 T32 MH19942-07). We thank William Vaughn for generating a pseudo-random sequence and Editor Dan Sanes and two anonymous reviewers for invaluable comments.

References

- Aertsen A, Johannesma P, Hermes D. Spectro-temporal receptive fields of auditory neurons in the grassfrog. II Biol Cybern. 1980; 38:235–248.
- Ahrens MB, Linden JF, Sahani M. Nonlinearities and contextual influences in auditory cortical responses modeled with multilinear spectrotemporal methods. J Neurosci. 2008; 28:1929–1942. [PubMed: 18287509]

- Andoni S, Li N, Pollak GD. Spectrotemporal receptive fields in the inferior colliculus revealing selectivity for spectral motion in conspecific vocalizations. *J Neurosci.* 2007; 27:4882–4893. [PubMed: 17475796]
- Atencio CA, Blake DT, Strata F, Cheung SW, Merzenich MM, Schreiner CE. Frequency-modulation encoding in the primary auditory cortex of the awake owl monkey. *J Neurophysiol.* 2007; 98:2182–2195. [PubMed: 17699695]
- Bar-Yosef O, Rotman Y, Nelken I. Responses of neurons in cat primary auditory cortex to bird chirps: effects of temporal and spectral context. *J Neurosci.* 2002; 22:8619–8632. [PubMed: 12351736]
- Barbour DL, Wang X. Auditory cortical responses elicited in awake primates by random spectrum stimuli. *J Neurosci.* 2003; 23:7194–7206. [PubMed: 12904480]
- Bodenhamer R, Pollak GD. Response characteristics of single units in the inferior colliculus of mustache bats to sinusoidally frequency modulated signals. *J Comp Physiol A.* 1983; 153:67–79.
- Brand A, Urban R, Grothe B. Duration tuning in the mouse auditory midbrain. *J Neurophysiol.* 2000; 84:1790–1799. [PubMed: 11024071]
- Brimijoin WO, O'Neill WE. On the prediction of sweep rate and directional selectivity for FM sounds from two-tone interactions in the inferior colliculus. *Hear Res.* 2005; 210:63–79. [PubMed: 16263230]
- Britt R, Starr A. Synaptic events and discharge patterns of cochlear nucleus cells. II. Frequency modulated tones. *J Neurophysiol.* 1976; 39:179–194. [PubMed: 1249601]
- Brosch M, Schreiner CE. Time course of forward masking tuning curves in cat primary auditory cortex. *J Neurophysiol.* 1997; 77:923–943. [PubMed: 9065859]
- Brosch M, Schreiner CE. Sequence sensitivity of neurons in cat primary auditory cortex. *Cerebral Cortex.* 2000; 10:1155–1167. [PubMed: 11073865]
- Calford MB, Semple MN. Monaural inhibition in cat auditory cortex. *J Neurophysiol.* 1995; 73:1876–1891. [PubMed: 7623087]
- Casseday JH, Ehrlich D, Covey E. Neural tuning for sound duration: Role of inhibitory mechanisms in the inferior colliculus. *Science.* 1994; 264:847–850. [PubMed: 8171341]
- Christianson GB, Sahani M, Linden JF. The consequences of response nonlinearities for interpretation of spectrotemporal receptive fields. *J Neurosci.* 2008; 28:446–455. [PubMed: 18184787]
- Covey E, Kauer JA, Casseday JH. Whole-cell patch-clamp recording reveals subthreshold sound-evoked postsynaptic currents in the inferior colliculus of awake bats. *J Neurosci.* 1996; 16:3009–3018. [PubMed: 8622130]
- de Bruijn NG. A Combinatorial Problem. *Koninklijke Nederlandse Akademie v Wetenschappen* 49. 1946; 49:758–764.
- deCharms RC, Blake DT, Merzenich MM. Optimizing sound features for cortical neurons. *Science.* 1998; 280:1439–1443. [PubMed: 9603734]
- Depireux DA, Simon JZ, Klein DJ, Shamma SA. Spectro-temporal response field characterization with dynamic ripples in ferret primary auditory cortex. *J Neurophysiol.* 2001; 85:1220–1234. [PubMed: 11247991]
- Eggermont JJ, Aertsen AM, Hermes DJ, Johannesma PI. Spectro-temporal characterization of auditory neurons: redundant or necessary. *Hear Res.* 1981; 5:109–121. [PubMed: 6976342]
- Escabi MA, Schreiner CE. Nonlinear spectrotemporal sound analysis by neurons in the auditory midbrain. *J Neurosci.* 2002; 22:4114–4131. [PubMed: 12019330]
- Faure PA, Fremouw T, Casseday JH, Covey E. Temporal masking reveals properties of sound-evoked inhibition in duration-tuned neurons of the inferior colliculus. *J Neurosci.* 2003; 23:3052–3065. [PubMed: 12684492]
- Fitzpatrick DC, Kanwal JS, Butman JA, Suga N. Combination-sensitive neurons in the primary auditory cortex of the mustached bat. *J Neurosci.* 1993; 13:931–940. [PubMed: 8441017]
- Fritz JB, Shamma SA, Elhilali M, Klein D. Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nat Neurosci.* 2003; 6:1216–1223. [PubMed: 14583754]
- Fritz JB, Elhilali M, Shamma SA. Differential Dynamic Plasticity of A1 Receptive Fields during Multiple Spectral Tasks. *J Neurosci.* 2005; 25:7623–7635. [PubMed: 16107649]

- Fuzessery ZM. Response selectivity for multiple dimensions of frequency sweeps in the pallid bat inferior colliculus. *J Neurophysiol.* 1994; 72:1061–1079. [PubMed: 7807196]
- Fuzessery ZM, Hall JC. Role of GABA in shaping frequency tuning and creating FM sweep selectivity in the inferior colliculus. *J Neurophysiol.* 1996; 76:1059–1073. [PubMed: 8871220]
- Fuzessery ZM, Richardson MD, Coburn MS. Neural Mechanisms Underlying Selectivity for the Rate and Direction of Frequency-Modulated Sweeps in the Inferior Colliculus of the Pallid Bat. *J Neurophysiol.* 2006; 96:1320–1336. [PubMed: 16790604]
- Gordon M, O'Neill WE. Temporal processing across frequency channels by FM selective auditory neurons can account for FM rate selectivity. *Hear Res.* 1998; 122:97–108. [PubMed: 9714578]
- Harris DM, Dallos P. Forward masking of auditory-nerve fiber responses. *J Neurophysiol.* 1979; 42:1083–1107. [PubMed: 479921]
- Heil P, Irvine DRF. Functional specialization in auditory cortex: responses to frequency-modulated stimuli in the cat's posterior auditory field. *J Neurophysiol.* 1998; 79:3041–3059. [PubMed: 9636107]
- Kaltenbach JA, Meleca RJ, Falzarano P. Forward masking properties of neurons in the dorsal cochlear nucleus. *Soc Neurosci Abstr.* 1992; 18:149.
- 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.* 1994; 96:1229–1254. [PubMed: 7962992]
- Klein DJ, Depireux DA, Simon JZ, Shamma SA. Robust spectrotemporal reverse correlation for the auditory system: optimizing stimulus design. *J Comput Neurosci.* 2000; 9:85–111. [PubMed: 10946994]
- Kowalski N, Depireux DA, Shamma SA. Analysis of dynamic spectra in ferret primary auditory cortex. I. Characteristics of single-unit responses to moving ripple spectra. *J Neurophysiol.* 1996; 76:3503–3523. [PubMed: 8930289]
- Kvale MN, Schreiner CE. Short-term adaptation of auditory receptive fields to dynamic stimuli. *J Neurophysiol.* 2004; 91:604–612. [PubMed: 14762146]
- Leroy SA, Wenstrup JJ. Spectral integration in the inferior colliculus of the mustached bat. *J Neurosci.* 2000; 20:8533–8541. [PubMed: 11069961]
- Linden JF, Liu RC, Sahani M, Schreiner CE, Merzenich MM. Spectrotemporal structure of receptive fields in areas AI and AAF of mouse auditory cortex. *J Neurophysiol.* 2003; 90:2660–2675. [PubMed: 12815016]
- Machens CK, Wehr MS, Zador AM. Linearity of cortical receptive fields measured with natural sounds. *J Neurosci.* 2004; 24:1089–1100. [PubMed: 14762127]
- Mendelson JR, Cynader MS. Sensitivity of cat primary auditory cortex (A1) neurons to the direction and rate of frequency modulation. *Brain Res.* 1985; 327:331–335. [PubMed: 3986511]
- Miller LM, Escabi MA, Read HL, Schreiner CE. Spectrotemporal receptive fields in the lemniscal auditory thalamus and cortex. *J Neurophysiol.* 2002; 87:516–527. [PubMed: 11784767]
- Mittmann DH, Wenstrup JJ. Combination-sensitive neurons in the inferior colliculus. *Hear Res.* 1995; 90:185–191. [PubMed: 8974996]
- Nataraj K, Wenstrup JJ. Roles of Inhibition in Complex Auditory Responses in the Inferior Colliculus: Inhibited Combination-Sensitive Neurons. *J Neurophysiol.* 2006; 95:2179–2192. [PubMed: 16371455]
- O'Neill WE, Frisina RD. Functional organization of mustached bat inferior colliculus I: Representation of FM2 frequencies carrying target range information. *Soc Neurosci Abstr.* 1985; 11:2151.
- O'Neill WE, Suga N. Target-range sensitive neurons in the auditory cortex of the mustached bat. *Science.* 1979; 203:69–73. [PubMed: 758681]
- Olsen JF, Suga N. Combination-sensitive neurons in the medial geniculate body of the mustached bat: Encoding of relative velocity information. *J Neurophysiol.* 1991; 65:1254–1274. [PubMed: 1875241]
- Peterson DC, Voytenko S, Gans D, Galazyuk A, Wenstrup JJ. Intracellular recordings from combination-sensitive neurons in the inferior colliculus. *J Neurophysiol.* 2008; 100:629–645. [PubMed: 18497365]

- Plack CJ, Oxenham AJ, Drga V. Masking by inaudible sounds and the linearity of temporal summation. *J Neurosci.* 2006; 26:8767–8773. [PubMed: 16928865]
- Portfors CV, Wenstrup JJ. Delay-tuned neurons in the inferior colliculus of the mustached bat: implications for analyses of target distance. *J Neurophysiol.* 1999; 82:1326–1338. [PubMed: 10482752]
- Portfors CV, Wenstrup JJ. Responses to combinations of tones in the nuclei of the lateral lemniscus. *J Assoc Res Otolaryngol.* 2001; 2:104–117. [PubMed: 11550521]
- Portfors CV, Wenstrup JJ. Excitatory and facilitatory frequency response areas in the inferior colliculus of the mustached bat. *Hear Res.* 2002; 168:131–138. [PubMed: 12117515]
- Rhode WS, Greenberg S. Lateral suppression and inhibition in the cochlear nucleus of the cat. *J Neurophysiol.* 1994; 71:493–514. [PubMed: 8176421]
- Rieke F, Bodnar DA, Bialek W. Naturalistic stimuli increase the rate and efficiency of information transmission by primary auditory afferents. *Proc Biol Sci.* 1995; 262:259–265. [PubMed: 8587884]
- Roverud RC. Neural computations for sound pattern recognition: evidence for summation of an array of frequency filters in an echolocating bat. *J Neurosci.* 1993; 13:2306–2312. [PubMed: 8501509]
- Rutkowski RG, Shackleton TM, Schnupp JW, Wallace MN, Palmer AR. Spectrotemporal receptive field properties of single units in the primary, dorsocaudal and ventrorostral auditory cortex of the guinea pig. *Audiol Neurotol.* 2002; 7:214–227.
- Sanchez JT, Gans D, Wenstrup JJ. Glycinergic “inhibition” mediates selective excitatory responses to combinations of sounds. *J Neurosci.* 2008; 28:80–90. [PubMed: 18171925]
- Schuller G, Radtke-Schuller S, Betz M. A stereotaxic method for small animals using experimentally determined reference profiles. *J Neurosci Methods.* 1986; 18:339–350. [PubMed: 3540473]
- Sen K, Theunissen FE, Doupe AJ. Feature analysis of natural sounds in the songbird auditory forebrain. *J Neurophysiol.* 2001; 86:1445–1458. [PubMed: 11535690]
- Shannon-Hartman S, Wong D, Maekawa M. Processing of pure-tone and FM stimuli in the auditory cortex of the FM bat, *Myotis lucifugus*. *Hear Res.* 1992; 61:179–188. [PubMed: 1326505]
- Shechter B, Depireux DA. Stability of spectro-temporal tuning over several seconds in primary auditory cortex of the awake ferret. *Neuroscience.* 2007; 148:806–814. [PubMed: 17693032]
- Singh NC, Theunissen FE. Modulation spectra of natural sounds and ethological theories of auditory processing. *J Acoust Soc Am.* 2003; 114:3394–3411. [PubMed: 14714819]
- Suga N. Analysis of frequency-modulated sounds by auditory neurons of echo-locating bats. *J Physiol.* 1965; 179:26–53. [PubMed: 5854378]
- Suga N, O'Neill WE. Neural axis representing target range in the auditory cortex of the mustached bat. *Science.* 1979; 206:351–353. [PubMed: 482944]
- Suga N, O'Neill WE, Kujirai K, Manabe T. Specificity of “combination sensitive” neurons for processing complex biosonar signals in the auditory cortex of the mustached bat. *J Neurophysiol.* 1983; 49:1573–1626. [PubMed: 6875639]
- Suga N, O'Neill WE, Manabe T. Cortical neurons sensitive to combinations of information-bearing elements of biosonar signals in the mustached bat. *Science.* 1978; 200:778–781. [PubMed: 644320]
- Suga N, Schlegel P. Coding and processing in the auditory systems of FM-signal-producing bats. *J Acoust Soc Am.* 1973; 54:174–190. [PubMed: 4731643]
- Suga N, Simmons JA, Shimozaawa T. Neurophysiological studies on echolocation systems in awake bats producing CF-FM orientation sounds. *J Exp Biol.* 1974; 61:379–399. [PubMed: 4443734]
- Sutter ML, Schreiner CE, McLean M, O'Connor KN, Loftus WC. Organization of inhibitory frequency receptive fields in cat primary auditory cortex. *J Neurophysiol.* 1999; 82:2358–2371. [PubMed: 10561411]
- Theunissen FE, Sen K, Doupe AJ. Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. *J Neurosci.* 2000; 20:2315–2331. [PubMed: 10704507]
- Theunissen FE, Woolley SM, Hsu A, Fremouw T. Methods for the analysis of auditory processing in the brain. *Ann N Y Acad Sci.* 2004; 1016:187–207. [PubMed: 15313776]

- Valentine PA, Eggermont JJ. Stimulus dependence of spectro-temporal receptive fields in cat primary auditory cortex. *Hear Res.* 2004; 196:119–133. [PubMed: 15464309]
- Walton JP, Frisina RD, Ison JR, O'Neill WE. Neural correlates of behavioral gap detection in the inferior colliculus of the young CBA mouse. *Journal of Comparative Physiology A, Sensory, Neural, and Behavioral Physiology.* 1997; 181:161–176.
- Warren S, Hamalainen HA, Gardner EP. Objective classification of motion- and direction-sensitive neurons in primary somatosensory cortex of awake monkeys. *J Neurophysiol.* 1986; 56:598–622. [PubMed: 3783213]
- Wehr M, Zador AM. Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex. *Nature.* 2003; 426:442–446. [PubMed: 14647382]
- Westerman LA, Smith RL. Rapid and short-term adaptation in auditory nerve responses. *Hear Res.* 1984; 15:249–260. [PubMed: 6501113]
- Wickesberg RE, Oertel D. Delayed, frequency-specific inhibition in the cochlear nuclei of mice: a mechanism for monaural echo suppression. *J Neurosci.* 1990; 10:1762–1768. [PubMed: 1972392]
- Woolley SM, Gill PR, Theunissen FE. Stimulus-Dependent Auditory Tuning Results in Synchronous Population Coding of Vocalizations in the Songbird Midbrain. *J Neurosci.* 2006; 26:2499–2512. [PubMed: 16510728]

Fig. 1.

Generation of the spectrotemporal response map (STRM): (A) FRA of a single unit showing a BEF of 52.0 kHz, MT of 18 dB SPL, and an excitatory bandwidth at MT+20 of 1.7 kHz. White squares indicate tone frequencies selected for use in the de Bruijn stimulus used to obtain the STRM. (B). Schematic of a small segment of the de Bruijn tone sequence centered at the BEF of the unit. Entire sequence was played 3-5 times and the resulting spike train was binned to create PST histograms using tone-triggered averaging, illustrated for two BEF tones in the sequence by the shaded areas. (C) A collection of 9 PST histograms corresponding to the 9 tone frequencies in the sequence, stacked in order of increasing frequency. (D) The single-tone STRM, a color-mapped version of the stacked histograms in

C. The hot spot indicates a strong response to tones at BEF with a latency of 8 ms. The dark region below BEF indicates low-sideband suppression, and the dark region following the BEF response indicates post-excitatory suppression. (E) Spike counts in response to the tone/FM/noise (TUDN) stimulus set. The unit responded strongly to tones (open circles, dotted line) and weakly to NB noise (open squares, dotted line) without duration selectivity. Response to up FM (filled green triangles) was poor at all durations. Response to down FM (filled red triangles) was weak for 32 ms duration and greater than that to tones at 16 and 8 ms. (F) Tone-pair STRMs, stacked by 1st tone frequency from low to high. Within each tone-pair STRM, histograms are stacked by 2nd tone frequency from low to high. The yellow spot in the top and bottom panels indicates a strong response to [+6 kHz → BEF] and [-6.0 kHz → BEF] pairs, respectively. The vertical band in the center panel indicates a good response to all [BEF → x] pairs. The hot spot in the +3.0 kHz panel indicates a facilitated response to BEF 2nd tones caused by +3.0 kHz 1st tones.

Fig. 2.

An example of *post-stimulatory suppression*: (A) FRA of a single unit showing a BEF of 58.2 kHz, MT of 20 dB SPL, and an excitatory bandwidth at MT+15 of 1.9 kHz. White squares indicate tone frequencies selected for use in the de Bruijn stimulus used to obtain the STRM. (B) Single-tone STRM. Sequence was centered at the unit's BEF, presented at MT +15 = 33 dB SPL. The hot spot indicates a response to tones from +1.5 to -1.5 kHz re BEF with a latency of 7 ms. The dark region following the BEF response indicates post-excitatory suppression. (C) Spike counts in response to the TUDN stimulus set. The unit responded strongly to 8 and 16 ms tones (open circles, dotted line) and 8 ms noise bursts (open squares, dotted line). Response to up FM (filled green triangles) was poor at all

durations. Response to down FM (filled red triangles) was weak for 32 ms duration and strong for 18 and 8 ms durations, matching that to tones. (D) Tone-pair STRMs. Some activity was found in response to 2nd tones at BEF for all 1st tones except those at BEF (middle panel). Suppression, seen as a notch in the vertical banding pattern is seen following all 2nd tones regardless of 1st tone frequency, and regardless of the magnitude of 2nd tone response. The red arrow indicates that post-excitatory suppression following 2nd BEF tones persists even when 2nd BEF tones did not drive any spiking activity.

Fig. 3.

An example of *frequency-specific suppression*: (A) FRA of a single unit showing a BEF of 47.0 kHz, MT of 12 dB SPL, and an excitatory bandwidth at MT+20 of 6.2 kHz. White squares indicate tone frequencies selected for use in the de Bruijn stimulus used to obtain the STRM. (B) Single-tone STRM. Sequence was centered at the unit's BEF of 58.8 kHz, presented at MT+20 = 35 dB SPL. The wedge of excitation indicates that the strong response at BEF (~5 ms latency) falls off in magnitude as tones move up in frequency. The dark region following the BEF response indicates post-excitatory suppression. (C) Spike counts in response to the TUDN stimulus set. The unit responded strongly to pure tones of all durations (open circles, dotted line) and weakly to noise bursts (open squares, dotted

line). Response to down FM (filled red triangles) was moderate at all durations. Response to up FM (filled green triangles) was matched that to tones except at 8 ms. (D) Tone-pair STRMs. A wedge of excitation similar to that seen in the single-tone STRM is seen in response to 2ⁿ tones when preceded by below-BEF 1st tones (bottom four panels). BEF tones suppress responses to all subsequent tones (center panel). Above-BEF 1st tones primarily suppress responses to above-BEF 2nd tones. Second tones at +1.5 kHz are suppressed by 1st tones at +1.5, but not -1.5, kHz (white arrows). Similarly, 2nd tones at +3.0 kHz are suppressed by 1st tones at +3.0, but not -3.0 kHz (red arrows). This frequency specificity results in a 2nd tone receptive field that is transiently altered by 1st tone frequency.

Fig. 4.

STRM predictions of FM directionality. (A) Single-tone STRMs: Convolution of the single-tone STRM was used to predict directional selectivity (DS) indices. These were compared with actual DS indices for 3 rates of FM. Perfect predictions would lie along the diagonal line. Filled squares represent 32 ms sweeps, open circles represent 16 ms sweeps, and crosses represent 8 ms sweeps. Actual DS indices ranged from -1.0 to +1.0, while the STRM predictions contained notably less dynamic range, with most values falling between -0.33 and +0.33 where directionality is considered non-significant. The linear single-tone STRM was unable to accurately capture the prevalent feature of FM directional selectivity. (B) Tone-pair STRMs: PSTHs from those tone-pairs whose effective modulation rate matched

that of the FM sweeps used in the TUDN stimulus were used to generate DS indices and compared with actual DS indices. Perfect predictions would lie along the diagonal line. For 32 ms FM sweeps (filled squares - linear regression: solid line), predicted DS indices were correlated with actual DS indices with an R^2 value of 0.47. For 16 ms FM sweeps (open circles - dashed line), this correlation was absent ($R^2 = 0.01$). For 8 ms FM sweeps (asterisks - dotted line), predicted DS indices were correlated with actual DS indices with an R^2 value of 0.22.

Fig. 5.

Conceptual Models: (A) A simple delayed sideband inhibition model used to explain basic FM directionality (left). Input from the high frequency excitatory branch (green) arrives prior to input from the lower-frequency inhibitory branch (red), allowing, in this example, down FM to drive excitation prior to the arrival of inhibition. Up FM would cause excitation to arrive simultaneously with inhibition, thus negating the response. A modified model (right) including an overlap in the tuning of the excitatory and inhibitory branches of the model, the tuning of each represented by overlapping dotted ovals. The delayed suppression (Δt) in this simple model explains the paradoxical post-excitatory suppression because the inhibitory arm is driven by the same range of tone frequencies that drive the excitatory arm.

(B) A two-stage delayed sideband inhibition model for frequency-specific suppression. The model consists of two independent excitatory inputs (output stage) from units 1 and 2 which may themselves be sideband- and on-BEF suppressed (this input stage is similar to the whole model seen in A). Depending on tone frequency and the tuning overlap of the input stage, arm 1 and/or arm 2 may be shut off due to sideband or post-excitatory suppression. In this example, high frequency tones suppress responses to subsequent high frequency tones by shutting off arm 2 of the model. BEF tones suppress all subsequent responses because they activate all four of the initial inputs. The sideband inhibition suggested by arm 1 is not seen in the example unit (Fig. 3), but is included here as a generalization to show the means by which low frequency tones could suppress responses to subsequent low-frequency tones.