Published in final edited form as:

Hear Res. 2008 May ; 239(1-2): 107–125. doi:10.1016/j.heares.2008.01.015.

Coding of FM sweep trains and twitter calls in area CM of marmoset auditory cortex

Yoshinao Kajikawa^{a,c}, Lisa A. de la Mothe^a, Suzanne Blumell^b, Susanne J. Sterbing-D'Angelo^a, William D'Angelo^a, Corrie R. Camalier^a, and Troy A. Hackett^{a,b}

^aDept. of Psychology, Vanderbilt University, Nashville, TN 37203

^bDept. of Hearing and Speech Sciences, Vanderbilt University School of Medicine, Nashville, TN 37203

^cNathan Kline Institute, Orangeburg NY

Keywords

temporal; synchrony; vocalization; primate; auditory

INTRODUCTION

Our working model of the primate auditory cortex includes approximately 12 areas distributed among three regions: core, belt, and parabelt (Kaas et al., 2000; Kaas et al., 1999). These areas are topographically interconnected, and each receives a unique blend of inputs from the medial geniculate complex (MGC) and several multisensory nuclei in the posterior thalamus (Hackett, 2007; Jones, 2007). The patterns of connections between areas suggest that auditory cortical processing involves serial and parallel elements (Kaas et al., 1998; Rauschecker et al., 1997). According to the model, information ascends a regional core-belt-parabelt hierarchy, representing three stages of processing. Within each region, areas work in parallel on inputs from several cortical and thalamic sources. Based on the hierarchical relationships between regions, neurons in later stages of processing (i.e., belt, parabelt) are expected to exhibit response properties distinct from that of earlier stages (e.g., core), including: (i) longer response latency; (ii) wider spectro-temporal integration, and (iii) greater stimulus selectivity (Rauschecker, 1998).

Recordings in core and belt areas of various primate species have produced a variety of evidence in support of the first two predictions. Neurons in the lateral belt areas generally have longer response latencies and wider spectro-temporal tuning compared to neurons in area A1 of the core region (Bieser et al., 1996; Kajikawa et al., 2005; Kosaki et al., 1997; Lakatos et al., 2005; Merzenich et al., 1973; Rauschecker et al., 2004; Rauschecker et al., 1995; Rauschecker et al., 1997; Recanzone et al., 2000a; Recanzone et al., 2000b; Tian et al., 2001; Woods et al., 2006). In addition, selectivity for the spatial and non-spatial features of sounds is greater in the belt region, and varies topographically between areas (Recanzone, 2000; Tian et al., 2001; Woods et al., 2006).

Corresponding author: Troy A. Hackett, Ph.D. Vanderbilt University 301 Wilson Hall 111 21st Avenue South Nashville, TN 37203
Phone: (615) 322-7491 Fax: (615) 343-8449 Email: troy.a.hackett@vanderbilt.edu.

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.

Given that neurons in the lateral belt tend to exhibit longer response latencies and increased spectro-temporal integration, a fourth prediction is that synchronization to repetitive or cyclic sounds would be degraded in the belt areas. In general, it has been observed that periodicity or temporal envelope coding degrades along the ascending auditory pathway (Joris et al., 2004), presumably due to the intervening synaptic relay (Berry et al., 1976; Miles, 1986; Sutor et al., 1989). If those principles also apply to auditory cortex, temporal resolution should be greater in the core than the belt areas. At present, almost nothing is known about temporal fidelity of neurons in the belt areas, as most studies have focused on neurons in A1 (Cheung et al., 2001; Liang et al., 2002; Luczak et al., 2004; Malone et al., 2007; Nagarajan et al., 2002; Wang et al., 2001; Wang et al., 1995). In a study of several areas in the awake squirrel monkey, Bieser and Muller-Preuss (1996) found that the average best modulation frequency (BMF) for sinusoidally amplitude-modulated tones (SAMt) in the lateral belt areas was lower than in A1, consistent with results obtained in A1 and AII of cats (Eggermont, 1998; Schreiner et al., 1986).

Although these results support predictions about areas in the core-belt-parabelt hierarchy, several studies suggest that at least one of the belt areas, the caudomedial belt area (CM), does not fully conform to the generic profile of a belt area. In ketamine-anesthetized marmosets, we reported that the distributions of minimum response latencies in A1 and CM were highly overlapping, but significantly shorter overall in CM for tones and noise bursts (Kajikawa et al., 2005). Tuning bandwidth was relatively broad in CM. In awake macaques, (Lakatos et al., 2005) found that response latencies in A1 for tones were shorter compared to CM, whereas latencies for noise were longer in A1. Finally, in the peri-insular belt area (Pi), which partly corresponds to CM, Bieser and Muller-Preuss (1996) reported shorter mean latencies for tones compared to A1, while BMFs for sinusoidal amplitude modulated tones were comparable among neurons in A1 and Pi (CM). These findings indicate that neurons in CM have broad spectral tuning typical of lateral belt areas, but also suggest that their temporal fidelity may be typical of neurons in A1.

In the present study, we set out to test this latter prediction in area CM of marmoset monkeys. Temporal coding was studied using two related classes of repetitive acoustic stimuli: 1) marmoset twitter call; and 2) frequency-modulated (FM) sweep trains. The marmoset twitter call is a multisyllabic vocalization, consisting of 6 to 8 upward FM phrases repeated at about 7 to 8 Hz. This call is frequently uttered during social interactions (Epplé, 1968), and has been the subject of several studies in A1 of marmosets (Cheung et al., 2005; Luczak et al., 2004; Nagarajan et al., 2002; Wang et al., 2001; Wang et al., 1995). Generally, neuron response profiles in A1 faithfully represent the twitter call's temporal envelope, with a preference for its natural form over temporally modulated or degraded versions of the call. The FM sweep trains were modeled after the phrases of the marmoset twitter call, but modified to estimate temporal fidelity across a range of repetition rates. The number of sweeps in a train was limited to 7, which approximates the average number of phrases in the twitter call. The temporal envelope was modulated by changing the repetition rate of successive sweeps within the train from 1 to 24 Hz, widely overlapping the typical envelope modulation frequency of the twitter call. The spectrotemporal structure of individual sweeps was held constant. Given the short response latency and broad spectral tuning observed in CM, we predicted that the temporal precision for twitter calls and FM sweeps would remain high across the 6 to 8 Hz range (natural phrase frequency), but would be degraded at higher rates.

The results of the present study confirmed this prediction, demonstrating that neurons in CM appear to contribute to the processing of vocalizations in a manner that preserves temporal details. This adds new insight into the functional significance of this auditory field and its place in the hierarchy. In addition, since auditory cortex is required for the discrimination of species-specific calls (Heffner et al., 1986; Heffner et al., 1989), information about their encoding by

individual fields contributes to our understanding of how complex sounds are processed within auditory cortex, in general.

MATERIALS and METHODS

Animal subjects

All experimental procedures were conducted in marmoset monkeys (*Callithrix jacchus jacchus*) in accordance with NIH Guidelines for the Use of Laboratory Animals under a protocol approved by the Vanderbilt University Institutional Animal Care and Use Committee. Multiunit recordings were obtained from auditory cortex in the right hemispheres of four ketamine-anesthetized marmoset monkeys, ages 2 to 5 years, during a terminal experiment, eight to ten days after neuroanatomical tracer injections were made into the left auditory cortex of each animal subject, as previously described (de la Mothe et al., 2006a). We limited recordings to the hemisphere opposite the injection site to minimize potential effects of tracer injections on neuron response properties. The experimental histories of these animals were detailed in a previous study in which responses to tones and noise bursts were described for most of the recording sites (Kajikawa et al., 2005). Preliminary findings were reported previously (Kajikawa et al., 2003).

General surgical procedures

After premedication with cefazolin (25 mg/kg), dexamethazone (2 mg/kg), cimetidine HCl (5 mg/kg), and robinul (0.015 mg/kg), anesthesia was induced by ketamine hydrochloride (10 mg/kg IM). A surgical plane of anesthesia was maintained by IV administration of ketamine, supplemented by IM injections of xylazine (0.4 mg/kg). Body temperature was kept at 37°C with a water circulating heating pad. Heart rate, expiratory CO₂, and O₂ saturation were continuously monitored throughout the surgery and used to adjust anesthetic depth. Oxygen was delivered passively through an endotracheal tube at a rate of 1 liter/minute.

The head was held by hollow ear bars affixed to a stereotaxic frame (David Kopf Instruments, Tujunga, CA). The skull was exposed by midline incision. The temporal muscle was retracted and a craniotomy was made to expose the right dorsal superior temporal gyrus, lateral sulcus, and overlying parietal cortex. After retraction of the dura mater, warm silicone was applied to the brain to prevent desiccation of the cortex. Photographs of the exposed cortical surface were taken for recording the locations of electrode penetrations in relation to blood vessels and the lateral sulcus.

Auditory stimulation

Experiments were conducted in a sound-proof chamber (Industrial Acoustics Corp., NY). Acoustic stimuli were generated by Tucker-Davis Technologies (TDT, Gainesville, FL) System II hardware and software (SigGen). Auditory stimulation was made by coupling STAX MK3 electrostatic earphones (Stax, Japan) to a hollow ear bar inserted into the ear contralateral to the recording sites in CM. Stimuli were calibrated using a 1/4" microphone (Model 7017; ACO Pacific, CA) and custom software (TDT, SigCal). Amplitude corrections were saved in a data file and applied to each stimulus to preequalize the response of the earphone. Stimulus presentation was controlled by the data acquisition software (see below).

After obtaining the frequency response area and intensity series of broad band noise (Kajikawa et al. 2005), trains of upward linear FM sweeps (30 ms duration; 5 ms cos² ramp; 72 dB SPL) were presented. These FM sweep parameters were modeled after the marmoset twitter call covering two different frequency ranges (1) FM₂₋₁₀ (2 to 10 kHz, 2.32 octaves/sec), (2) FM₆₋₁₄ (6 to 14 kHz, 1.22 octaves/sec). The acoustic parameters of the latter are comparable to the natural twitter call of the marmoset. The former covers a lower spectral range, and was

used because it corresponds more closely to the response range of low CF neurons. During a stimulus acquisition period of 3.5 seconds, trains of identical FM sweeps were presented at 11 different repetition rates (4 sweeps at 1 Hz, 7 sweeps each from 2 to 20 Hz in 2 Hz steps). At each rate, the FM train was presented 10 times (110 total trials). These stimuli were chosen because we sought to examine temporal fidelity more systematically using repetitive sounds that modeled the twitter call. This vocalization is usually composed of a series of phrases for which power, spectral structure and intervals between phrases vary within the same call. Since temporal modulation by phase vocoding also alters power of each phrase, we used FM sweep trains over a range of repetition rates to investigate temporal coding fidelity. One benefit of these stimuli is that the temporal envelope can be manipulated parametrically while keeping the energy of each phrase constant and maintaining a rough similarity of the twitter call sound structure. In addition, whereas tEMF may change within a single twitter call, the repetition rate of the FM sweep trains was predetermined, and the envelope frequency of the response to be examined was known.

In most experiments, digitized marmoset twitter calls (courtesy of X. Wang, Johns Hopkins) were presented after the FM sweep trains. Each of the three exemplars (tw A, tw B, tw C) had peak temporal modulation frequencies (tEMF) of 6.5, 8.5, and 7.5 Hz, respectively. These calls were selected to better represent the acoustic variability in the calls made by individuals in this species. For a subset of 8 units, responses were obtained from modified calls, generated from tw A or B by temporal compression of the spectrogram to one half of their original duration with a phase vocoder technique (Flanagan and Golden, 1966; Wang et al., 1995). These calls, labeled half tw A and B, had tEMF peaks of 13.5 and 17 Hz. The tEMF of half tw A was not precisely twice that of the original call due to frequency aliasing. The frequency interval of the vector strength (VS) spectrum was 0.5 Hz. All calls were presented 15 times at 72 dB RMS.

Data acquisition

Multiunit (MU) recordings were made with 1 M Ω tungsten microelectrodes insulated with Parylene C or polyimide (MicroProbe Inc., Potomac, MD). Electrodes were advanced using a hydraulic or piezoelectric microdrive (David Kopf Instruments, Tujunga, CA). To target CM on the lower bank of the LS, vertical penetrations were made through the overlying parietal cortex. Because of the angle of the LS, the electrode angle was slightly oblique relative to the pial surface. At each recording site, clusters of neurons were identified in layer III between about 500 and 900 μ m below the pia. Depth was tracked relative to the transition between the upper and lower banks of the LS. The locations of recording sites were identified in reconstructed serial sections of each brain using tonotopic gradients (Kajikawa et al., 2005) and recently established cytoarchitectonic criteria (de la Mothe et al., 2006a).

Stimulus presentation and data acquisition were controlled by Brainware software (TDT, Tucker-Davis Technologies, Gainesville, FL) and TDT System II hardware. In most electrode penetrations, single channel recordings were made, but in some penetrations, recordings were obtained simultaneously from an array of two to four electrodes at approximately the same depth separated by 0.5 or 1 mm. Signals were amplified (10,000x) and filtered (300 to 3000 Hz). Waveforms (500 or 1000 ms re stimulus onset) and timestamps of spikes were saved on a computer hard drive for offline sorting and analysis. At the end of each recording procedure, electrolytic marker lesions (10 μ A for 10 seconds) were made in key locations to facilitate reconstruction of recording sites.

Data analysis

Phase-locked firing rate—For each of 7 FM sweeps in a train, we estimated the average firing rate in the period from 0 – 30 ms from sweep onset (duration of individual sweeps). Since there were 7 sweeps in a train (except for 4 sweeps at 1 Hz repetition rate), there were

7 (or 4) estimates of the firing rate for each trial. The mean of 10 trials was plotted for each sweep to represent the entrainment pattern in the manner similar to a previous study (Eggermont, 1991). The rate modulation transfer function (rMTF) was obtained by further averaging the mean firing rates of the last 6 (3) sweeps in a train and plotting the estimate for the repetition rate. The best modulation frequency (rBMF) was defined as the repetition rate at which the rMTF peaked.

The response during the 30 ms period included most of the phase-locked firing to FM sweeps (referred to below as “phase-locked transients”). However, firing rate was estimated during this period regardless of whether there were phase-locked transients. In the absence of phase-locked transients, the entrainment pattern represents the shape of tonic activity.

We also performed LDA of the entrainment pattern. Each pattern was expressed as a vector of 7 components, equal to the number of sweeps in a train. For 1 Hz, the train contained only 4 sweeps; therefore, firing rates for the fifth to seventh sweeps were randomly resampled from those of the first to fourth sweeps, since the 1 Hz response was steady in general. LDA was applied to those vectors as described below for classification of PSTHs, but without dimension reduction.

Vector strength—Synchronization to the FM sweep train at each repetition rate was quantified for all spikes occurring after the onset of the second sweep until the end of the stimulation period by vector-strength ($VS = \sum_t e^{2\pi i t/T}$, t = spike times, T = cycle length, i = imaginary unit). Since the stimulus repetition rate was known before analysis, the cycle length for VS estimation was fixed to the one derived from the repetition rate for each response. VS quantifies the magnitude of regularity in synchronization or phase-locking to cyclic stimulation. The temporal modulation function (vsMTF) was estimated from the relationship between the repetition rate of sweeps and VS. The significance of VS was determined using a Rayleigh test ($P < 0.01$). Some trials contained multiple response peaks within a single repetition cycle. Those peaks tended to cancel each other, reducing VS, and often led to non-significant synchronization by the Rayleigh test. Therefore, the synchrony of such complex temporal patterns was further characterized by a novel statistical test (shuffle D test) of temporal structure or spike train synchrony based on entropy analysis (Kajikawa and Hackett, 2005). Thus, in some instances, we regarded the synchrony of complex temporal response patterns as significant according to the shuffle D test ($P < 0.05$), even if the Rayleigh test was not significant. The repetition rate at which the function of significant VS peaked was selected as the vsBMF. The highest repetition rate at which VS was significant was identified as maxF, which was equivalent to the term ‘synchronization-boundary’ in some studies (Lu et al., 2001).

Quantification of synchrony in response to twitter calls—The PSTH for each twitter call was constructed from responses recorded over 15 trials. The distribution of response power in the frequency domain was calculated for each response pattern. To make comparisons of peaks from different spectra (e.g. response to tw A and half tw A) possible, the dimensions of spectra were scaled to that of the VS by:

$$VS(f) = \sqrt{F(PSTH) \bullet F^*(PSTH) / \sum_t PSTH},$$

in which $F(\cdot)$ denotes Fourier transform:

$$F(x(t)) = F(f | x(t)) = \sum_t \{x(t) \bullet e^{2\pi i f t}\},$$

and $F^*(\cdot)$ denotes its complex conjugate. Thus the square root of the complex conjugate-squared PSTH (equivalent to the auto-power spectrum) was taken and then normalized by the first order power of itself. To detect significant temporal modulation of PSTH at different

frequencies, a similar VS-spectrum was estimated for 1000 PSTHs, whose inter-spike interval sequences were shuffled from the original sequence. This destroyed the structure of the temporal pattern, but preserved the number of spikes. Peaks and troughs of the VS-spectrum of the original spike train that were below or above the 2.5 to 97.5 percentiles of the shuffled VS-spectrum were considered significant ($P < 0.05$). The magnitudes of peaks across different spectra could be compared because spectra were expressed in dimensions of the VS, which compensated for the difference in the firing rates. Unlike the response to the FM sweep train, we did not analyze the response to twitter call using VS at a single frequency, because intervals between the onsets of phrases were not fixed within individual calls (DiMattina et al., 2006). Dynamic changes in the temporal envelope modulation frequency (tEMF) also made pre-determination of tEMF ambiguous.

Classification of temporal response patterns—To examine the capacity of the temporal firing pattern to predict the repetition rate or twitter call exemplar identity, we performed Fisher's linear discriminant analysis (LDA). This analysis classifies the target responses based on the known structures of reference response patterns. To suppress data bias in classification, LDA included cross-validation (Romanski et al., 2004). For the response to twitter calls, one out of 15 trials for each call was randomly extracted and 3 trials from all calls were reserved as test trials. All remaining 42 trials were assigned as reference trials whose structures were used to classify 3 test trials. Similarly, for responses to FM sweep trains between 1 and 20 Hz, one out of 10 trials at each repetition rate (i.e., 11 out of 110 total trials), was randomly extracted as a test trial for classification. The remaining 99 trials were assigned as reference trials used to classify the 11 test trials. The peri-stimulus time histogram (PSTH) was obtained for each of the 11 separated trials.

In this analysis, the number of variables was limited by the number of reference trials (42 or 99 for twitter calls and FM sweeps, respectively). Accordingly, the dimension of the PSTH (number of bins) was reduced by principal component analysis (PCA). This yielded dimensions of 300 for twitter calls (1500 ms period / 5 ms bin-size and 325 for FM sweep trains (2600 ms period / 8 ms bin-size). After PCA, PCs were sorted by corresponding eigenvalues in descending order from the largest. We estimated the minimum number of eigenvalues whose summation exceeded 99% of the total variance in the original PSTH. PCs with a corresponding eigenvalue less than the number of trials used for classification were used for further analysis with LDA. To confirm that the PSTH generated from the reduced PCs contained most of the original structure, we estimated the cross correlation between the PSTH recovered from linear summation of those PCs and the original PSTH. Correlation coefficients ranged from 0.91 to 0.99 for twitter calls, 0.89 to 0.99 for FM sweep trains, and over 0.92 for 95% of the analyzed patterns, indicating that most of structure in the PSTH was well represented by PC reduction. Among those PCs, LDA was performed according to the algorithm described in the Appendix.

By repeating the procedure described above 1000 times, extraction-derived bias was reduced. Error or success of classification was counted at each repetition rate (or twitter call) and divided by 1000 to estimate correct or incorrect classification rates creating the confusion matrix. The average of all classification rates over repetition rates (or twitter calls) comprised the grand mean error rate (MER). For FM sweeps trains, we also identified the repetition rate with the least error as the best classification repetition rate (BCF). BCF represents the repetition rate at which the response pattern was most discernible or classified correctly.

One strength of using LDA is that it uses all structural information of the response, irrespective of pattern (Romanski et al., 2005). Thus, the response pattern was utilized without degrading into certain response parameters and discarding others. While techniques other than LDA have also been used for classification, it has been shown that LDA is better for decoding complex temporal patterns (Middlebrooks et al., 1998; Nicolelis et al., 1998). Other methods available

for classification of neuronal response patterns, like the brute force approach, require large samples of data (Petersen et al., 2002). In the present study, this sampling problem was overcome by LDA. Practically, the number of trials and length of stimulation limit the temporal resolution of this analysis. However, as long as the dominant features of the response pattern are represented in the PSTH, the analysis can be performed properly.

From the confusion matrix obtained for each MU from LDA, we estimated mutual information (MI) and mean error rate. MI was estimated by $MI = \sum_s \{p(s) \cdot \sum_r p(r|s) \log_2(p(r|s)/p(r))\}$, in which s and r represents stimulated and classified repetition rate respectively. $p(s)$ or $p(r)$ represents the probability of the corresponding repetition rate regardless of r or s . $P(r|s)$ is the conditional probability of classifying the repetition rate of the response as r , given that of stimulation s . Since our experiment had 11 repetition rates, the maximum possible value of MI in the case of perfect classification would be $\log_2(11) = 3.46$. Since analysis of the entrainment pattern also provides a similar confusion matrix, we estimated MI for both temporal response pattern and entrainment pattern.

RESULTS

Multiunits recorded from CM

This study includes data recorded from 42 multi-units (MUs) in area CM of 4 animals. All were examined with 3 twitter vocal sounds, and 8 MUs were also examined with temporally compressed versions of 2 twitters. Twenty three (55%) were examined with both FM₂₋₁₀ and FM₆₋₁₄, 2 (5%) were examined with only FM₂₋₁₀, and the remaining 17 (40%) were examined with only FM₆₋₁₄. More MUs were examined with FM₆₋₁₄ because its temporal structure and spectrum was modeled after the twitter calls and consequently prioritized during data collection. MUs with latencies longer than 30 msec and slow rates of rise in firing rate were omitted from further analysis (3 of 25, FM₂₋₁₀; 5 of 40, FM₆₋₁₄). Consequently, 22 and 35 MU responses to FM₂₋₁₀ and FM₆₋₁₄, respectively, were described below.

A previous study examined the response of the same MUs to tones and noise (Kajikawa et al., 2005). Fig. 1 summarizes the responses to tones measured at an intensity equal to that of the FM sweep trains and twitter vocal sounds (72 dB SPL). Bandwidth is indicated by gray bars for each MU described in the present study. Other than broadly-tuned units that responded across the entire frequency range at threshold, CFs tended to cluster into low and high frequency regions, and center frequency weighted by the magnitude of response to tones at 72 dB SPL tended to shift toward higher or lower frequencies. Tuning bandwidth of MUs in CM increased as sound intensity increased, and covered the frequency range of FM₂₋₁₀ and FM₆₋₁₄.

Responses to FM sweep trains

In the analyses of FM sweep train response patterns, our particular interest was to determine the optimal response parameters for coding the envelope modulation frequency range around 7–8 Hz, which corresponds most closely to the natural phrase frequency of the marmoset twitter call. Most of the MUs in CM exhibited an entrained response, but the response patterns were variable. In general, the magnitude of phase-locked transients decreased at higher repetition rates as a more slowly changing tonic firing pattern emerged. First, we describe three examples that exemplify the patterns observed, then move on to analyses of the sample.

FM sweep train responses: individual examples—The response of one MU to the FM₂₋₁₀ sweep train is shown in Fig. 2. This MU responded with only phase-locked transients at all repetition rates, without a change in the background firing between such peaks. This pattern is apparent in the raster plot and PSTH (Figs. 2A and B, respectively). While responses to each sweep were similar at 1Hz, the phase-locked responses to later sweeps at other rates

were larger than responses to the first sweeps, indicating enhancement at most repetition rates. Plotted in Fig. 2C are the phase-locked transient responses normalized to the firing rate of the first sweep in the train. At repetition rates above 4 Hz, the responses to sweeps 2 through 6 exceeded one, indicating enhancement of the response for those sweeps. Synchronization, reflected by *vsMTF*, was close to one at most repetition rates (black circles in Fig. 2D). Synchronization to FM sweeps was significant at all repetition rates. At 1 and 2 Hz, VS was reduced by temporally scattered background activity between phase-locked transients and smaller phase-locked responses.

Figs. 3 and 4 show the responses of two different MUs to the FM₆₋₁₄ sweep train. Both MUs showed degraded phase-locking at repetition rates above 10 Hz, but in different ways. The MU in Fig. 3 responded with phase-locked transients at low repetition rates. As repetition rate increased, phase-locked transients to successive sweeps diminished, converting gradually to a seemingly random, tonic firing pattern (Fig. 3A). PSTHs at repetition rates from 14 to 24 Hz were also flat after the phase-locked transient to the first sweep (Fig. 3B). Entrainment did not increase, and was slightly depressed at high repetition rates (Fig. 3C). Since phase-locked transients diminished, both *vsMTF* and *rMTF* decreased gradually as the repetition rate increased (Fig. 3D). Synchronization was not significant at repetition rates higher than 14 Hz.

The MU in Fig. 4 exhibited complete suppression of firing activity for about 100 ms after the phase-locked transient following the first sweep at all repetition rates (Fig. 4A and B). The suppression was followed by strong rebound tonic activity at higher repetition rates (> 10 Hz). Nonetheless, weak phase-locked transients were superimposed upon the tonic response, and synchronization was significant at all repetition rates (Fig. 4D). At 1 or 2 Hz, prominent rebound activity arose at about 200 ms after the first sweep (4A and B), reducing VS at those rates (4D). As a consequence of these features, entrainment was enhanced for successive sweeps at low repetition rates; whereas, at higher repetition rates, entrainment was reduced for the second sweep, and subsequently enhanced for later sweeps (Fig. 4C). Such complexity was also reflected in the *rMTF*, which peaked at 4 Hz and decreased gradually as the repetition rate increased (gray line in Fig. 4D), typical of most units in our sample (Fig. 5).

FM sweep train responses: population data—Among the response patterns observed in this study, variability was mainly reflected in the degree to which phase-locking was maintained at different repetition rates, and could therefore be evaluated by VS. There were also other aspects of variability. For example, phase-locking was modified at different repetition rates by loss or depression of phase-locked transients, the emergence of increased background/tonic firing rate, or both. In addition, tonic firing could appear as either a random flat or dynamically changing pattern of different strengths. Below, we describe those features in terms of the following parameters: (1) *rMTF*: How were the responses to FM sweep trains accompanied by changes in the mean magnitude of phase-locked transients? (2) *vsMTF*: How was synchronization of MU responses to FM sweep trains altered by the periodicity of sound? (3) Repetition rate classification by response pattern: How discernible was a change in the response pattern for different repetition rates or sound envelopes, by accounting for the dynamic structure of the response?

Rate modulation transfer functions

For most MUs, the response strength of phase-locked transients decreased as repetition rate increased (e.g. Fig. 3). The resulting mean trajectory of the *rMTF* was a monotonically decreasing function of the repetition rate (Fig. 5). The large SD was attributable to a few MUs for which the response was sustained or facilitated, even at high repetition rates (e.g. Fig. 2), or others that exhibited strong rebound from suppression after the first sweep. The latter enhanced the magnitude of phase-locked transients at repetition rates below 10 Hz (e.g. Fig.

- 4). The rBMF was above 10 Hz for both FM₂₋₁₀ and FM₆₋₁₄ in less than 15% of the MUs of our sample.

Synchronization modulation transfer function

Phase-locking of the response to FM sweep trains was quantified by VS. Fig. 6A shows the vsMTF population means. The mean vsMTF peak was 2 and 4 Hz for FM₂₋₁₀ and FM₆₋₁₄, respectively. The decline of vsMTF at 1 and 2 Hz reflected two response patterns: 1) Phase-locked transient onset, followed by rebound activity at those repetition rates (Fig. 4); or 2) Augmentation of phase-locked transient magnitude at higher repetition rates (Fig. 2). At rates above 2 Hz, VS generally decreased as the repetition rate increased. For most MUs, this was again due to the depression of phase-locked transients or the emergence of tonic firing (e.g., Fig. 3). In accordance with the mean vsMTF peak, median vsBMFs were 2 and 4 Hz for FM₂₋₁₀ and FM₆₋₁₄, respectively (Fig. 6B). The broad vsMTF error bars suggested variable diminution of phase-locking. For example, SDs at high repetition rates were greater for MUs with highly synchronized responses through 24 Hz (e.g., Fig. 2), despite a general decline of VS. Accordingly, the distribution of vsBMF was skewed toward slightly higher repetition rates (mean vsBMF: 4.95 Hz and 5.09 Hz for FM₂₋₁₀ and FM₆₋₁₄ respectively), and the values of maxF reached a maximum at 24 Hz (the highest rate tested). For FM₂₋₁₀ and FM₆₋₁₄, 44% (10/22) and 34% (12/35) of MUs could follow up to 20 Hz, with maxF means of 18.2 Hz and 15.0 Hz, respectively (Fig. 6C). Since 24 Hz was the highest repetition rate examined in the present study, we cannot state whether some MUs may synchronize to higher rates.

Mean response latencies at 1 Hz were 11.85 ms (SD = 1.32) for FM₂₋₁₀ and 12.89 ms (SD = 3.15) for FM₆₋₁₄. There was a significant negative correlation between maxF and response latency to FM sweeps (Spearman rank correlation one-tail test: $r_s = -0.443$, $P < 0.025$ and $r_s = -0.608$, $P < 0.0005$ for FM₂₋₁₀ and FM₆₋₁₄, respectively). Similar relationships between latency and maxF were previously described in area A1 of marmosets for sinusoidal AM and FM sounds (Liang et al., 2002).

Classification of the repetition rate by temporal response pattern

Evaluation of neuronal responses in terms of mean firing rate and phase-locking is assumed to reflect stimulus encoding by CM neurons. These analyses would have greater validity if neuronal activity was time-invariant; however, response patterns contained a complex mixture of phase-locked transients and tonic activity that fluctuated between cycles. Although this dynamic structure may be useful in decoding information contained in the response, those elements were either filtered out by mean firing rate or collapsed by VS. Therefore, we used LDA to determine how the dynamic features of the response pattern may contribute to classification of repetition rate.

First, similarity (or ambiguity) between PSTH pairs at different repetition rates was estimated by computing the cross-correlation coefficient. Examples of similarity for three MUs are shown in panel E of Figs. 2 - 4. The MUs in Figs. 3E and 4E had a region of lightly colored pixels at the top right corner representing PSTH pairs at repetition rates from 10 - 20 Hz, indicating that pairs at those rates were grossly similar compared to other pairs. Presumably, this is due to elevated tonic activity at higher repetition rates, since relatively high similarity would interfere with discrimination and cause errors of classification. Otherwise, there was no prominent trend in similarity between pairs at high repetition rates (compare Fig. 2E with Fig. 3E and 4E). In contrast, similarity was minimal at repetition rates below 12 Hz, where response patterns were dominated by phase-locked transients (e.g., Fig. 2). This was due, in part, to matching of the phase-locked transients produced by the FM sweep at 500 ms that were common to FM trains at repetition rates from 2-12 Hz.

Next, in panel F of Figs. 2 to 4, the results of LDA-based classification of FM sweep repetition rate by temporal response pattern were presented in a confusion matrix format similar to panel E for three MUs (from Figs. 2-4). Classification error rate at each repetition rate is shown in panel G. Classification error from 2–10 Hz was low for all three MUs. For the MUs in Figs. 3 and 4, classification error rate was high from 12–20 Hz, consistent with greater similarity between PSTH pairs (i.e., panel E). Classification of the response at 1 Hz was frequently erroneous as well, because spontaneous activity was either not suppressed between phase-locked transients, or because of similarities in tonic activity. For example, classification at 1 Hz was worse for the MU in Fig. 3 than Fig. 4, because weak tonic activity after the phasic response to the first sweep resembled the tonic response to successive sweeps at repetition rates above 12 Hz. In contrast, for the MU in Fig. 4, prominent late rebound after the first sweep at 1 Hz was distinct from the immediate rebound from suppression at high repetition rates. Accordingly, classification of repetition rate could be low cut (Fig. 2), low pass (Fig. 4), or band-pass (Fig. 3) within the range of repetition rates examined.

The LDA classification results are summarized in Fig. 7. Despite broad SDs, curves of the mean classification error rate for FM_{2–10} and FM_{6–14} exhibited common trends (Fig. 7A): (1) Classification by temporal response pattern had a band-pass profile, where the lowest mean classification error rate was observed at 4 Hz and remained less than 0.5 between 2–8 Hz; (2) Error rate was relatively high at 1 Hz and increased between 8–20 Hz; (3) Error rate reached a plateau at repetition rates above 14 Hz. Spontaneous or weak tonic activity between phase-locked transients contributed to the increase in error at low repetition rates, as in Figs. 2 and 3. At high repetition rates, error was increased by either weak tonic activity from diminished phase-locking (Fig. 3) or strong rebound activity that overwhelmed the phase-locked response (Fig. 4). A band of good classification was observed in the 4–6 Hz range, except for MUs that had poor synchronization even at 2 or 4 Hz (e.g., Fig. 6C). The distribution of BCF (Fig. 7B) peaked in this range for both FM_{2–10} (mean 6.21 Hz) and FM_{6–14} (mean 5.07 Hz), consistent with the curves in Fig. 7A.

Mean BCF values were similar to vsBMF, and there were no significant differences between vsBMF and BCF for either FM_{2–10} or FM_{6–14} (Wilcoxon signed rank test, $P > 0.2$). The result of this comparison suggested that information coded by phase-locking accounts for most of the information coded in the temporal response pattern; therefore, in addition to estimating synchrony, decoding of the temporal response pattern was redundant. However, there was no significant correlation between BCF and vsBMF for either FM type (Fig. 8A, Spearman's rank correlation, $P > 0.1$), suggesting independence of the information captured by those analyses. The nature of that disparity was revealed by plotting the difference between BCF and vsBMF (i.e., BCF - vsBMF) against vsBMF (Fig. 8B). This comparison revealed a significant correlation for both FM types ($P < 0.001$). BCF was higher than vsBMF when vsBMF was below 6 Hz, and vice versa. Thus, BCF tended to be around 6 Hz, regardless of vsBMF, presumably reflecting the distinctness of the temporal response pattern near this rate.

Classification of the repetition rate by entrainment pattern

The entrainment pattern produced by phased-locked firing varied with repetition rate, suggesting another potential strategy for coding of the repetition rate. At lower repetition rates, for example, entrainment grew gradually from a flat pattern (Figs. 2C–4C, lighter gray lines), whereas the pattern varied among MUs at higher rates (darker lines). However, classification on the basis of entrainment pattern was significantly worse than for temporal response patterns. The median MER for temporal pattern was 0.62 and 0.64, whereas the median MER for entrainment pattern was 0.82 and 0.83 for FM_{2–10} and FM_{6–14} respectively. (Wilcoxon signed rank test, $P < 0.0001$ for both FM types). Accordingly, mutual information was significantly lower for entrainment pattern than for temporal response pattern in the classification of

repetition rate. The median MI for temporal response pattern was 1.01 and 0.92, while the median MI for entrainment pattern was 0.67 and 0.54 for FM₂₋₁₀ and FM₆₋₁₄ respectively (Wilcoxon signed rank test, $P < 0.005$ for both FM types). The reason could be that the difference in timing of those sweeps was ignored in the firing rate profiles. Thus, it is likely that classification by mean firing rate is relatively poor because the dynamics of the phasic firing rate among sweeps are lost.

Response to twitter calls in CM

The marmoset twitter call is usually composed of repeating short phrases of 30 – 40 ms with a tEMF near 8 Hz (Fig. 9A) (DiMattina et al., 2006). A sonogram of one exemplar is shown in Fig. 9A, along with the associated PSTHs of four CM neurons (Fig. 9B). MUs in CM typically responded robustly to twitter calls in an envelope-following manner such that the tEMF of the response peaked at around the tEMF of the call. MUs with different patterns of response to FM can entrain to the twitter call. Examples of the response to tw A are shown in Fig. 9B (panels a – d). Panels a – c show the responses of 3 MUs for which periodic peaks followed each phrase of the call. The response tEMF deviated from the call tEMF by no more than ± 1 Hz, and VS compared to shuffled spike trains was significant for all three calls ($P < 0.05$), as shown in Fig. 9C (panels a-c). Allowing for a ± 1 Hz deviation, 55.0 % (22/40) of the MUs in this study responded to all 3 calls with significant peaks at the tEMF, 17.5 % (7/40) responded significantly to 2 calls, and 15.0 % responded to only 1 of 3 calls. Of 40 MUs, periodicity at the tEMF was significant for 29 MUs to tw A, 27 for tw B, and 30 for tw C. MUs for which the response was not significantly periodic had either no response at all or a tonic rise in firing rate with no entrainment. An example is shown in Fig. 9B (panel d) for tw A, with its VS profile in Fig. 2C (panel d).

Eight MUs with significantly periodic responses to both tw A and B ($P = 0$) were further examined with half tw A and B. Fig. 10 shows an example of a response to tw B and half tw B. The peak of the VS spectrum was significant for 6 MUs at the tEMF of both calls ($P < 0.05$). While the 95 % confidence intervals of the two calls overlapped, there was a significant trend for VS magnitude at the tEMF to be larger for tw A or B than half tw A or B (Wilcoxon signed rank test, $P < 0.05$). These results suggested that responses to faster tEMFs were less entrained.

The spectral analyses described so far are based on Fourier analysis and are thus most appropriate for data which display stationarity (Bendat et al., 2000), however, twitter call responses are constrained in time, contain syllabic response components with variable energy and temporal patterns, and are accompanied by variable tonic activity. To incorporate such response profiles into our analyses, we expressed each response as a vector of spike counts per time bin, which reflects any response profile. We performed LDA of those vectors in a cross-validation procedure, in which the response pattern was classified into a stimulation condition based on knowledge of a separate group of reference response patterns recorded under all stimulus conditions. If the response pattern to a particular stimulus was persistent on every trial and distinct from responses to other stimuli, it could be discriminated from the others and correctly classified into the actual stimulus condition. Conversely, if the response patterns to different stimuli were similar, ambiguity would make classification incorrect. From a functional point of view, this analysis judges whether neural responses are capable of the discrimination between different calls.

Illustrated in Fig. 11 are examples of 3 MU responses in CM to each of the 3 calls. Envelope synchrony for the MU in Fig. 11Ba was higher compared to the other 2 MUs, and its response pattern classified the 3 calls without error (Fig. 11Ca). In contrast, the response of the MU in Fig. 11Bb did not follow the sound envelope well and its classification error was relatively large (Fig. 11Cb). The response of one MU with low VS and classification error for all 3 calls is shown in Fig. 11Bc. Fig. 11Cc reveals that while entrainment was similar to the MU in Fig.

11Bb, presumably due to continually high tonic activity during sound, call classification by response pattern was much better and comparable to the MU in Fig. 11Ba.

Classification error rates did not differ significantly among the 3 twitter calls: tw A, 0.30 (SD: 0.30); tw B, 0.30 (SD: 0.26); and tw C, 0.27 (SD: 0.28) (Kruskal-Wallis, $P > 0.71$). Rather, classification depended on the temporal patterns of their responses. For each call, the classification error of responses that were entrained at the call's tEMF, as determined by the significance of the VS spectral peak, was significantly less than that of MUs without entrainment (Mann-Whitney U test, $P < 0.05$). In Fig. 12, classification error rate is plotted against the VS peak at the tEMF of the call for all MU responses with significant tEMF entrainment. As a whole, there was a significant negative correlation between classification error rate and the VS peak (Spearman rank correlation: $r_s = -0.453$, $P < 0.0005$). Thus, even for MUs with significant entrainment, there was a trend toward better sound envelope classification with better response synchrony. However, there were also MUs with weaker entrainment for which classification error was low (e.g., points at the lower left corner of Fig. 12). Therefore, the coding of twitter calls may be characterized not solely by the entrainment to the sound envelope, but also additional features in the response pattern.

Relationship between responses to FM sweep trains and twitter calls

Finally, we examined the relationships between responses to the FM sweep trains and twitter calls. Since the spectrotemporal properties of FM sweep trains in the 6 – 8 Hz range were similar to natural twitter calls, it was predicted that responses to those sounds would be correlated. First, the peak VS spectral magnitude of the response to twitter calls near their tEMFs (6.5, 8.5 and 7 Hz for tw A, B, and C, respectively) was compared with VS in response to the FM sweep train at the repetition rate nearest the tEMF of the twitter call (Table 1, VS). The VS peak of twitter call response tended to correlate positively with the VS of the FM sweep train response, suggesting that envelope following by MUs in our sample was comparable for both classes of stimuli. The correlation between classification error rates for twitter calls and FM sweep trains at the tEMF of the call (Table 1, Error rate) was also robust, indicating that the temporal response pattern significantly contributed to envelope classification of these sounds. Taken together, these results suggest that the temporal envelope of the stimulus is reflected in the response of CM neurons, and provides a basis for reliable classification.

DISCUSSION

In the present study, we analyzed multiunit responses in area CM of the marmoset monkey auditory cortex to twitter calls and trains of FM sweeps modeled after those calls. The main findings of the present study were:

1. Neurons in area CM responded in a phase-locked manner to the temporal envelope of twitter calls and trains of FM sweeps over the range of the twitter call phrase frequency
2. LDA revealed that information about the repetitive sounds used in this study was encoded not only by phase-locking, but also other components of the temporal response pattern.
3. CM entrains to repetitive sounds, and the pattern was similar to A1 (see below)

Overall, these results indicate that neurons in CM are capable of temporal processing at least as rapid as the vocalization sound envelope. As for A1, neurons in CM may also participate in processing that requires relatively precise temporal envelope discrimination, suggesting that CM is positioned at an early stage in the auditory processing hierarchy of primates.

Synchronization to repetitive sounds in CM

Synchronization is perhaps the most informative feature of the neuronal response to the repetitive elements of an auditory stimulus. As an index of such temporal coding, the calculation of VS has been useful in revealing differences among neurons, such as progressive reductions in phase-locking by nuclei in the ascending auditory pathway (Joris et al., 2004), or the presence of a duality in repetition rate coding in A1 (Liang et al., 2002; Lu et al., 2004; Lu et al., 2001). In the present study, we used VS to analyze multiunit responses in area CM of the marmoset monkey auditory cortex to multi-syllabic twitter calls and trains of FM sweeps modeled after those calls. Most of the MUs in our sample responded in a phase-locked manner to the temporal envelope of these stimuli over the behaviorally-relevant range of the twitter call. For the FM sweep trains, synchrony was best between 2 and 8 Hz, overlapping the range of the twitter call phrase frequency (~7 – 8 Hz), and decreased systematically through 24 Hz.

The reduction of vsMTF at higher repetition rates was attributable to dynamic processes, such as depression of phase-locked transients or changes in the rate of tonic activity. In general, VS does not capture information about such dynamic changes in synchronization due to the way it is calculated (i.e., by averaging the extent of phase-locking across cycles). The same limitation applies to mean firing rate, where ambiguity can result from the reduced capability to resolve the high dimensional characteristics of the stimulus. For example, firing rates depend on modulation frequency as well as intensity (Bartlett et al., 2005; Eggermont, 1994b; Malone et al., 2007; Phillips et al., 1989a; Phillips et al., 1989b). Equivalent firing rates could be attained by a continuum of combined repetition rates and sound intensities, therefore repetition rate cannot reliably be discriminated, suggesting that resolving the temporal response pattern may facilitate decoding.

Discrimination of repetitive sounds by response patterns in CM

Generally, the coding of repetitive sounds by temporal response patterns has not been adequately addressed in the literature, even for A1. While it appears that cortical neurons use synchronized or non-synchronized strategies to code cyclic sound (Lu et al., 2001), additional dimensions are possible. For example, the response pattern can diverge considerably for naturalistic or more complex sounds (Bar-Yosef et al., 2002; Newman et al., 1973; Wollberg et al., 1972). A single sound can evoke various response patterns among neurons, and a single neuron can respond differentially to many sounds. Further, a neuronal response may not necessarily follow the sound envelope. Even so, analyses of the response pattern can be used to discriminate complex sounds using an approach such as LDA (Romanski et al., 2005). Several studies have applied similar response pattern analyses to reveal coding parameters for vocalizations (Nelken et al., 2005; Romanski et al., 2005; Sugihara et al., 2006) or sound localization (Furukawa et al., 2001; Nelken et al., 2005; Stecker et al., 2003a; Stecker et al., 2005).

In the present study, we employed LDA to decode repetition rate from the response pattern as expressed in the PSTH. Response classification by LDA revealed that while phase-locking to the dominant periodic elements of twitter calls and FM sweep trains accounted for much of the information coded in the temporal response pattern, other components in the response pattern contributed to coding, as well. Our results also indicated that while VS was reduced by certain aspects of the neuronal response (e.g., post-onset suppression), LDA utilized such features to discriminate the sound envelope. Interestingly, the best discrimination by temporal response pattern did not correlate with the best synchronization, suggesting that envelope coding was not necessarily limited to response entrainment, but utilized additional aspects of the response pattern. In particular, synchronization to the FM sweep train correlated better with synchronization to twitter calls than with response discrimination. Therefore, repetitive sounds

may be coded by a template-like response pattern in addition to synchrony, suggesting that additional information can be extracted from the neuronal response by incorporating pattern analysis along with analyses of firing rate and synchronization.

Response synchrony in auditory cortex: comparisons with previous studies

Responses to cyclic auditory stimuli of various types have been described in many studies of auditory cortex of awake or anesthetized animals, but mainly for A1. Although comparisons between studies are limited by methodological variations (e.g., the type of stimuli used, species, cortical area, and anesthetic state), analyses have typically utilized VS to obtain indices such as temporal best modulation frequency (tBMF) and the upper frequency limit of synchronization (synchronization boundary), which are equivalent to vsBMF and maxF in the present study.

In A1 of barbiturate-anesthetized marmosets, entrainment to the temporal envelope was precise for the natural twitter call, but synchrony was disrupted when the temporal envelope was degraded or compressed (Wang et al., 1995; Nagarajan et al., 2002). When Nagarajan et al. (2002) compared the responses to twitter calls, sAM tones at CF, and click trains, the strength of synchronization to calls correlated with that of sAM tones and click trains, but only weakly. We found similar results in CM. Entrainment was the most precise for natural calls, and a weak correlation for synchronization magnitude was observed between twitter calls and FM sweep trains.

Malone et al (2007) examined responses to SAM tones at modulation frequencies below 20 Hz. Their analysis of the modulation period histogram revealed that all of the parameters of the SAM stimulus (i.e., carrier frequency and level, modulation frequency and depth) were reflected in the response. Thus, the MTF alone does not adequately characterize the response to such stimuli. Those findings complement the main finding of the present study, in that information relevant to discrimination is contained in the synchronous and non-synchronous aspects of the response pattern.

To our knowledge, the only data of this kind available for area CM were collected in awake squirrel monkeys (Bieser et al., 1996). In that study, synchronization to sAM tones was studied in A1, R and several areas of the belt region. In both A1 and area Pi, which corresponds to the medial portion of CM, tBMF ranged from 8 – 16 Hz. For neurons in area Pa, which corresponds to the remainder of CM and part of CL, tBMF ranged from 2 – 8 Hz. In the lateral belt areas, tBMFs were generally lower, ranging from 1 – 8 Hz. Despite difference in anesthetic state, vsBMFs for CM neurons in the present study ranged from 2 – 8 Hz with a mean of about 5 Hz, and were therefore comparable to neurons in Pa of the awake squirrel monkey.

Our results also fall into the range of tBMFs obtained in A1 and AAF of anesthetized cats. Mean tBMFs for trains of either tones or clicks ranged from about 5 to 8 Hz, consistent with the relatively short response latencies typically recorded in these areas. (Eggermont, 1991; Eggermont, 1993; Eggermont, 1998; Eggermont et al., 1995; Phillips et al., 1989b; Schreiner et al., 1996). Interestingly, mean values of tBMF in anesthetized cats and rats are substantially higher for sinusoidally modulated AM or FM tones or noise than trains of tones or noise (Eggermont, 1991; Eggermont, 1993; Gaese et al., 1995; Kilgard et al., 1999; Orduna et al., 2001; Schreiner et al., 1988; Schreiner et al., 1996). These findings indicate that tBMFs tend to be lower for stimulus trains than sinusoidally-modulated stimuli, suggesting that trains of repetitive short-duration stimuli, with multiple onsets and offsets, engage fundamentally different processes in auditory circuits than continuously modulated stimuli, which have a single onset event.

With respect to the upper limit of synchronization, or synchronization boundary, the mean maxF measured in this study was between 15 and 20 Hz. The maxF values of several MUs reached the highest repetition rate examined (24 Hz), suggesting that the upper limit may have been higher for those units. Comparisons with other studies are generally limited to results from A1, and are made even more difficult by rather large variability. In A1 of anesthetized cats, average synchronization boundaries in different studies have ranged from 40 Hz for sAM (Eggermont, 1998), 25 Hz or less for click trains (Lu et al., 2000), to less than 7 Hz for tone pulse trains (Phillips et al., 1989b). In A1 of anesthetized rats values of maxF have typically fallen below 18 Hz for sAM/sFM and trains of tones or FM sweeps (Creutzfeldt et al., 1980; Gaese et al., 1995; Kilgard et al., 1999; Orduna et al., 2001). In A1 of awake animals, synchronization to click trains have been reported to reach 100 Hz or higher in some studies (de Ribaupierre et al., 1972; Goldstein et al., 1959; Lu et al., 2001; Steinschneider et al., 1998), whereas in other studies, maximum synchronization for sFM/sAM was less than 20 Hz (Bieser et al., 1996; Liang et al., 2002; Schulze et al., 1999; Whitfield et al., 1965). In general, it appears that mean maxF for CM neurons in anesthetized marmosets is comparable to that observed in A1 of several studies.

Suppression of response synchrony for repetitive sounds

In comparing our results to studies of A1 and AAF, we noticed that synchronization as indexed by tBMF and maxF is commonly affected by two factors. The first is a period of suppression after the onset response, which is followed by a rebound from suppression in a subsequent cycle (Aitkin et al., 1968; Creutzfeldt et al., 1980; Fitzpatrick et al., 1999; Lu et al., 2000; Nelken et al., 2000). This is most important for repetitive sounds with a total duration that overlaps the rebound period. In several studies, for example, total durations were 1 second or longer (Bieser et al., 1996; Liang et al., 2002; Lu et al., 2001; Orduna et al., 2001; Phillips et al., 1989b; Schreiner et al., 1986; Schreiner et al., 1996), or included repetition rates only down to 10 Hz (Creutzfeldt et al., 1980; Lu et al., 2001). In the present study, the number of repeated cycles was limited to 7, thus the total duration of cyclic stimulation was relatively short at repetition rates above 8 Hz compared to other studies. The second factor is depression of the phase-locked transient, especially for repetitive sounds with shorter cycle lengths (Creutzfeldt et al., 1980; Eggermont, 1991; Eggermont, 1994a; Kilgard et al., 1999; Lu et al., 2001). Their consistency over a broad range of experimental conditions, indicates that suppression is a general feature of the response to temporally modulated stimuli, as well as for sounds with durations that overlap the time-scale of the suppression.

Interpretation of modulation transfer functions: functional considerations

Firing rate and VS, as shown by the rMTF and vsMTF functions, are dependent on the repetition rate, and while they certainly constitute information-bearing elements of the neuronal response, they do not fully reflect temporal response patterns. These measures can be interpreted as how neurons represent stimulation parameters physiologically (Lu et al., 2000; Lu et al., 2001). Then, the stimulus that evoked the highest magnitude of the response parameter, (e.g., rBMF or vsBMF) is considered as the best stimulation parameter represented by the neuron.

However, a different picture emerges from the view that those neurons contribute to discrimination of sound location or vocal call type, as suggested by auditory cortical lesion studies (Heffner et al., 1986; Heffner et al., 1990; Jenkins et al., 1984). Theoretically, the greatest sensitivity of an encoder to a certain range of stimulation can be achieved by aligning the dynamic range of the encoder response to the variable region of the events to be coded (Dean et al., 2005; Laughlin, 1981). Such alignment increases the information transmission between stimulation and response. Recent studies elucidating the coding of stimulus parameters suggested that neurons may code the critical range of stimulation not by the magnitude of the response, but by its gradient (McAlpine et al., 2001; Stecker et al., 2005).

For example, the steepest response gradient to the interaural time difference exists near the midline where the ability of animals to discriminate sound location is the greatest.

In the present study, even though the FM sweep trains were modeled from twitter call, which have an envelope modulation frequency of 7~8 Hz, the range of repetition rates identified as the best by rMTF and vsMTF were different. In contrast, for either rMTF or vsMTF, it was notable that the steepest slope of the mean function was between 4 and 10 Hz, which encompasses the modulation frequency range of twitter call (Figs. 2 and 3). According to the functional view described above, this may mean that CM neurons discriminated envelope modulation frequencies within the range of 4–10 Hz better than repetition rates above and below that range. Above 4–10 Hz, repetition rate was best represented by firing rate magnitude, whereas synchronization was most important below this range. Furthermore, since the range of steepest monotonic change was similar for rMTF and vsMTF, the redundancy may combine to enhance the reliability of the coding of that range. Although such a possibility could be explored using techniques like the maximum-likelihood procedure, which is similar to LDA, our preliminary test suggested that 10 trials were not enough to obtain probability distributions of firing rate or VS for each repetition rate. It is interesting to note that the 4–10 Hz range also corresponds to the range of repetition rates that were discriminated well by temporal response patterns. Thus, we speculate that both firing rate and synchronization probably contributed to discrimination by the temporal response pattern. In addition, these observations suggest that even though rBMF or tBMF may represent physiological properties of neurons, a discriminative function may be attributed to the range of stimulation parameters over which the response varied most. Since the 4–10 Hz range may be behaviorally important, better discrimination within this range might take on greater functional significance in CM.

The role of CM in auditory cortical processing

For the most part, anatomical and physiological studies of the primate auditory cortex have emphasized the role of areas within the core region, a variety of new information about the belt and parabelt regions has accumulated within the last decade. Beginning with Rauschecker et al (1997), area CM has been the focus of several studies. In that study, it was reported that after removal of A1 by aspiration, neurons in CM became unresponsive to tones, but remained responsive to complex sounds, suggesting that activity in CM was at least partly dependent on a serial relay from A1. That study was followed by a series of anatomical studies which revealed that the caudal and rostral areas of auditory cortex were connected with functionally-distinct areas of parietal and frontal cortex (Hackett et al., 1999; Romanski et al., 1999a; Romanski et al., 1999b). This topography supported the hypothesis that spatial (where) and non-spatial (what) auditory processing were mediated by partly segregated processing streams with origins in the caudal and rostral areas of auditory cortex, respectively (Kaas et al., 1999; Kaas et al., 2000; Rauschecker et al., 2000). Physiological support for this hypothesis came from several studies in which sensitivity to spatial features was observed in the caudal belt, including CM (Recanzone, 2000; Recanzone et al., 2004; Recanzone et al., 2000b; Tian et al., 2001; Woods et al., 2006). Additional insight into the role of CM came from the discovery of multisensory integration in this area (Fu et al., 2003; Schroeder et al., 2001). A subpopulation of neurons in CM were found to be responsive to both auditory and somatosensory stimulation. Adjacent areas of somatosensory cortex (e.g., retroinsular area, Ri) appear to be the source of that somatosensory input (de la Mothe et al., 2006a; Hackett, 2007; Smiley et al., 2007). Other studies have shown that CM is tonotopically organized, and that response latencies are as short as, or shorter than, neurons in A1 (Bieser et al., 1996; Kajikawa et al., 2005; Lakatos et al., 2005). A potential anatomical explanation for these features is that the principal thalamic input to CM derives from the rostral pole of the MGC, and perhaps MGv, as well (de la Mothe et al., 2006b; Hackett et al., 2007). In cats, neurons in the rostral pole, which has strong connections to AAF, are tonotopically organized and have short response latencies comparable

to MGv (Imig et al., 1984; Imig et al., 1985; Lee et al., 2005; Lee et al., 2004; Morel et al., 1987). Finally, the results of the present study indicate that neurons in CM are capable of temporal processing at least as rapid as the vocalization sound envelope. Since response latencies and responses to vocalization are similar in A1, it appears that neurons in A1 and CM process several features of the auditory scene in parallel. Thus, CM could be engaged in a wide range of activities that require relatively precise temporal envelope discrimination, suggesting that CM is positioned at an early stage in the auditory processing hierarchy of primates.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Xiaoqin Wang for helpful and insightful comments on previous versions of this manuscript, and for use of the twitter calls. We also thank the reviewers for excellent recommendations and feedback.

GRANTS

This research was supported by NIH/NIDCD R01 04318 to T.A.H.

Appendix

The LDA algorithm used in the present study was based on Fisher's original LDA formulation (*below*, vectors and matrices are in bold type). Data were recorded under R stimulation conditions equal to the number of repetition rates, subscripted by r , each symbolized by π_r ($r = 1 \cdots R$). Each condition had n_r trials (10 or 15 in the present study) and there were a total of N trials ($N = \sum_r^R n_r$). For each trial, the PSTH obtained from the spike train was expressed by a column vector, $\mathbf{x}_i^r$ (indices i and r stand for i th trial under a stimulation condition, π_r), of p components (p is the number of observed variables, equal to the number of bins in PSTH or the number of PCs after PCA). For cross-validation, n_r trials were separated from each stimulation condition into one as a test trial, expressed by $\mathbf{z}_r$, and the rest ($n_r - 1$ trials) as classifier trials. Within each condition, the mean of classifier trial observations was $\boldsymbol{\mu} = \sum_i^{n_r-1} \mathbf{x}_i^r / (n_r - 1)$. The grand mean of classifier trials over all stimulation conditions was $\boldsymbol{\mu}_0 = (1/(N-R)) \cdot \sum_r^R (n_r - 1) \cdot \boldsymbol{\mu}_r$. In total, there were R test trials (expressed by $\mathbf{x}'$ below) and $N-R$ classifier trials. Then, Fischer's discriminant function was applied. This was expressed as a linear combination of $Y_r = \mathbf{w} \cdot (\mathbf{z} - \boldsymbol{\mu}_r)$, which gives discriminant score Y_r , a scalar value representing the extent to which the test trial $\mathbf{z}$ is close to $\boldsymbol{\mu}_r$ in p dimension. Accordingly, given $Y_{j,s}$ for a particular test trial $\mathbf{z}_j$, this trial is classified to a condition π_s if $Y_{j,s}$ calculated with mean observation $\boldsymbol{\mu}_s$ gives the minimum among $Y_{j,s}$.

Derivation of linear weight vector $\mathbf{w}$ uses only classifier trials, as described below (Lachenbruch, 1998). The aim of the derivation is to make the variance of the discriminant score between π s larger than the variance within each π with regard to only classifier trials. First, among classifier trials, we obtain between and within conditions sum of covariance matrices of observations as $\mathbf{S}_B = \sum_r^R (\boldsymbol{\mu}_r - \boldsymbol{\mu}_0) \cdot (\boldsymbol{\mu}_r - \boldsymbol{\mu}_0)^T$ and $\mathbf{S}_w = \sum_r^R \sum_i^{n_r-1} (\mathbf{x}_i^r - \boldsymbol{\mu}_r) \cdot (\mathbf{x}_i^r - \boldsymbol{\mu}_r)^T$, respectively. Then, concerning the discriminant score of classifier trials with arbitrary $\mathbf{w}$, the ratio of variance between conditions to variance within conditions can be expressed by $(\mathbf{w}^T \mathbf{S}_B \mathbf{w}) / (\mathbf{w}^T \mathbf{S}_w \mathbf{w})$. We ought to maximize this ratio, and this can be achieved by using eigenvectors of $\mathbf{S}_w^{-1} \cdot \mathbf{S}_B$. To estimate its inverse matrix, $\mathbf{S}_w$ must be regular. This restricts p , the number of variables or components in every vector, to a value smaller than $N-R$, the number of classifier trials, to avoid singularity. From $\mathbf{S}_w^{-1} \cdot \mathbf{S}_B$, we calculated $R-1$ eigenvectors $\mathbf{w}_\lambda$ whose corresponding eigenvalues are top $R-1$ of p eigenvalues when sorted in descending order. Finally, a linear weight vector to calculate the discriminant score was obtained from $\mathbf{w} = \sum_\lambda^{R-1} \mathbf{w}_\lambda$.

LIST OF ABBREVIATIONS

A1, Primary (core) auditory area 1
 AII, Auditory area II
 AAF, Anterior auditory field
 BMF, Best modulation frequency
 BW, Bandwidth
 CF, Characteristic frequency
 CL, Caudolateral belt area
 CM, Caudomedial belt area
 FM, Frequency-modulated
 FRA, Frequency response area
 LDA, Linear discriminant analysis
 LS, Lateral sulcus
 maxF, Maximum frequency of synchronization
 MGC, Medial geniculate complex
 MGv, Ventral division of the MGC
 MGd, Dorsal division of the MGC
 MU, Multiunit
 MTF, Modulation transfer function
 PCA, Principal component analysis
 rBMF, Rate best modulation frequency
 rMTF, Rate modulation transfer function
 VS, Vector strength
 vsMTF, Vector strength modulation transfer function

List of Abbreviations

A1, Auditory area I (core)
 AII, Auditory area II
 AAF, Anterior auditory field
 AM, Amplitude modulated
 BCF, Best classification frequency
 BMF, Best modulation frequency
 CM, Caudomedial belt area (medial belt)
 IM, Intramuscular
 FM, Frequency modulated
 LDA, Linear discriminant analysis
 LS, Lateral sulcus
 maxF, Maximum synchronization frequency
 MER, Mean error rate
 MGC, Medial geniculate complex
 MGv, Medial geniculate complex, ventral division
 MI, Mutual information
 MU, Multiunit
 PCA, Principal component analysis
 rBMF, Rate best modulation frequency
 rMTF, Rate modulate transfer function
 SAMt, Sinusoidally amplitude modulated tones
 tEMF, Temporal envelope modulate frequency
 tw, Twitter call
 VS, Vector strength
 vsBMF, Vector strength best modulation frequency

vsMTF, Vector strength modulation transfer function

REFERENCES

- Aitkin LM, Dunlop CW. Interplay of excitation and inhibition in the cat medial geniculate body. *J Neurophysiol* 1968;31:44–61. [PubMed: 5640739]
- 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;22:8619–32. [PubMed: 12351736]
- Bartlett EL, Wang X. Long-lasting modulation by stimulus context in primate auditory cortex. *J Neurophysiol* 2005;94:83–104. [PubMed: 15772236]
- Bendat, JS.; Piesol, AG. *Random Data Analysis and Measurement Procedure*. Wiley; New York: 2000.
- Berry MS, Pentreath VW. Criteria for distinguishing between monosynaptic and polysynaptic transmission. *Brain Res* 1976;105:1–20. [PubMed: 175886]
- Bieser A, Muller-Preuss P. Auditory responsive cortex in the squirrel monkey: neural responses to amplitude-modulated sounds. *Exp Brain Res* 1996;108:273–84. [PubMed: 8815035]
- Cheung SW, Bedenbaugh PH, Nagarajan SS, Schreiner CE. Functional organization of squirrel monkey primary auditory cortex: responses to pure tones. *J Neurophysiol* 2001;85:1732–49. [PubMed: 11287495]
- Cheung SW, Nagarajan SS, Schreiner CE, Bedenbaugh PH, Wong A. Plasticity in primary auditory cortex of monkeys with altered vocal production. *J Neurosci* 2005;25:2490–503. [PubMed: 15758157]
- Creutzfeldt O, Hellweg FC, Schreiner C. Thalamocortical transformation of responses to complex auditory stimuli. *Exp Brain Res* 1980;39:87–104. [PubMed: 6247179]
- de la Mothe LA, Blumell S, Kajikawa Y, Hackett TA. Cortical connections of auditory cortex in marmoset monkeys: core and medial belt regions. *J Comp Neurol* 2006a;496:27–71. [PubMed: 16528722]
- de la Mothe LA, Blumell S, Kajikawa Y, Hackett TA. Thalamic connections of auditory cortex in marmoset monkeys: core and medial belt regions. *J Comp Neurol* 2006b;496:72–96. [PubMed: 16528728]
- de Ribaupierre F, Goldstein MH Jr. Yeni-Komshian G. Cortical coding of repetitive acoustic pulses. *Brain Res* 1972;48:205–25. [PubMed: 4645205]
- Dean I, Harper NS, McAlpine D. Neural population coding of sound level adapts to stimulus statistics. *Nat Neurosci* 2005;8:1684–9. [PubMed: 16286934]
- DiMattina C, Wang X. Virtual vocalization stimuli for investigating neural representations of species-specific vocalizations. *J Neurophysiol* 2006;95:1244–62. [PubMed: 16207780]
- Eggermont JJ. Rate and synchronization measures of periodicity coding in cat primary auditory cortex. *Hear Res* 1991;56:153–67. [PubMed: 1769910]
- Eggermont JJ. Differential effects of age on click-rate and amplitude modulation-frequency coding in primary auditory cortex of the cat. *Hear Res* 1993;65:175–92. [PubMed: 8458750]
- Eggermont JJ. Neural interaction in cat primary auditory cortex II. Effects of sound stimulation. *J Neurophysiol* 1994a;71:246–70. [PubMed: 8158231]
- Eggermont JJ. Temporal modulation transfer functions for AM and FM stimuli in cat auditory cortex. Effects of carrier type, modulating waveform and intensity. *Hear Res* 1994b;74:51–66. [PubMed: 8040099]
- Eggermont JJ. Representation of spectral and temporal sound features in three cortical fields of the cat. Similarities outweigh differences. *J Neurophysiol* 1998;80:2743–64. [PubMed: 9819278]
- Eggermont JJ, Smith GM. Rate covariance dominates spontaneous cortical unit-pair correlograms. *Neuroreport* 1995;6:2125–8. [PubMed: 8595185]
- Epple G. Comparative studies on vocalization in marmoset monkeys (Hapalidae). *Folia Primatol* 1968;8:1–40. [PubMed: 4966050]
- Fitzpatrick DC, Kuwada S, Kim DO, Parham K, Batra R. Responses of neurons to click-pairs as simulated echoes: auditory nerve to auditory cortex. *J Acoust Soc Am* 1999;106:3460–72. [PubMed: 10615686]
- Fu KM, Johnston TA, Shah AS, Arnold L, Smiley J, Hackett TA, Garraghty PE, Schroeder CE. Auditory cortical neurons respond to somatosensory stimulation. *J Neurosci* 2003;23:7510–5. [PubMed: 12930789]

- Furukawa S, Middlebrooks JC. Sensitivity of auditory cortical neurons to locations of signals and competing noise sources. *J Neurophysiol* 2001;86:226–40. [PubMed: 11431504]
- Gaese BH, Ostwald J. Temporal coding of amplitude and frequency modulation in the rat auditory cortex. *Eur J Neurosci* 1995;7:438–50. [PubMed: 7773441]
- Goldstein MH, Kiang NY-S, Brown RM. Responses of the auditory cortex to repetitive acoustic stimuli. *J Acoust Soc Am* 1959;31:356–364.
- Hackett TA, Stepniewska I, Kaas JH. Prefrontal connections of the parabelt auditory cortex in macaque monkeys. *Brain Res* 1999;817:45–58. [PubMed: 9889315]
- Hackett TA, De La Mothe LA, Ulbert I, Karmos G, Smiley J, Schroeder CE. Multisensory convergence in auditory cortex, II. Thalamocortical connections of the caudal superior temporal plane. *J Comp Neurol* 2007;502:924–52. [PubMed: 17444488]
- Hackett TA, Smiley JF, Ulbert I, Karmos G, Lakatos P, de la Mothe LA, Schroeder CE. Sources of somatosensory input to the caudal belt areas of auditory cortex. *Perception* 2007;36:1419–1430. [PubMed: 18265825]
- Heffner HE, Heffner HS. Effect of unilateral and bilateral auditory cortex lesions on the discrimination of vocalizations by Japanese macaques. *J Neurophysiol* 1986;56:683–701. [PubMed: 3783215]
- Heffner HE, Heffner RS. Cortical deafness cannot account for the inability of Japanese macaques to discriminate species-specific vocalizations. *Brain Lang* 1989;36:275–85. [PubMed: 2920286]
- Heffner HE, Heffner RS. Effect of bilateral auditory cortex lesions on sound localization in Japanese macaques. *J Neurophysiol* 1990;64:915–31. [PubMed: 2230934]
- Imig TJ, Morel A. Topographic and cytoarchitectonic organization of thalamic neurons related to their targets in low-, middle-, and high-frequency representations in cat auditory cortex. *J Comp Neurol* 1984;227:511–39. [PubMed: 6470221]
- Imig TJ, Morel A. Tonotopic organization in lateral part of posterior group of thalamic nuclei in the cat. *J Neurophysiol* 1985;53:836–51. [PubMed: 3981239]
- Jenkins WM, Merzenich MM. Role of cat primary auditory cortex for sound-localization behavior. *J Neurophysiol* 1984;52:819–47. [PubMed: 6512590]
- Jones, EG. *The Thalamus*. 2nd ed.. Cambridge University Press; Cambridge: 2007.
- Joris PX, Schreiner CE, Rees A. Neural processing of amplitude-modulated sounds. *Physiol Rev* 2004;84:541–77. [PubMed: 15044682]
- Kaas JH, Hackett TA. ‘What’ and ‘where’ processing in auditory cortex. *Nat Neurosci* 1999;2:1045–7. [PubMed: 10570476]
- Kaas JH, Hackett TA. Subdivisions of auditory cortex and processing streams in primates. *Proc Natl Acad Sci U S A* 2000;97:11793–9. [PubMed: 11050211]
- Kajikawa Y, Hackett TA. Entropy analysis of neuronal spike train synchrony. *J Neurosci Methods* 2005;149:90–3. [PubMed: 16026849]
- Kajikawa Y, de la Mothe LA, Blumell S, Hackett TA. Response properties of neurons in core and medial belt auditory cortex of marmoset monkeys. *Assoc. Res. Otolaryngol. Abs* 2003;972.
- Kajikawa Y, de la Mothe LA, Blumell S, Hackett TA. A comparison of neuron response properties in areas A1 and CM of the marmoset monkey auditory cortex: tones and broad band noise. *J Neurophysiol* 2005;93:22–34. [PubMed: 15342713]
- Kilgard MP, Merzenich MM. Distributed representation of spectral and temporal information in rat primary auditory cortex. *Hear Res* 1999;134:16–28. [PubMed: 10452372]
- Kosaki H, Hashikawa T, He J, Jones EG. Tonotopic organization of auditory cortical fields delineated by parvalbumin immunoreactivity in macaque monkeys. *J Comp Neurol* 1997;386:304–16. [PubMed: 9295154]
- Lakatos P, Pincze Z, Fu KG, Javitt DC, Karmos G, Schroeder CE. Timing of pure tone and noise-evoked responses in macaque auditory cortex. *Neuroreport* 2005;16:933–937. [PubMed: 15931064]
- Laughlin S. A simple coding procedure enhances a neuron's information capacity. *Z Naturforsch [C]* 1981;36:910–2.
- Lee CC, Winer JA. Principles governing auditory cortex connections. *Cereb Cortex* 2005;15:1804–14. [PubMed: 15800026]

- Lee CC, Imaizumi K, Schreiner CE, Winer JA. Concurrent tonotopic processing streams in auditory cortex. *Cereb Cortex* 2004;14:441–51. [PubMed: 15028648]
- Liang L, Lu T, Wang X. Neural representations of sinusoidal amplitude and frequency modulations in the primary auditory cortex of awake primates. *J Neurophysiol* 2002;87:2237–61. [PubMed: 11976364]
- Lu T, Wang X. Temporal discharge patterns evoked by rapid sequences of wide- and narrowband clicks in the primary auditory cortex of cat. *J Neurophysiol* 2000;84:236–46. [PubMed: 10899199]
- Lu T, Wang X. Information content of auditory cortical responses to time-varying acoustic stimuli. *J Neurophysiol* 2004;91:301–13. [PubMed: 14523081]
- Lu T, Liang L, Wang X. Temporal and rate representations of time-varying signals in the auditory cortex of awake primates. *Nat Neurosci* 2001;4:1131–8. [PubMed: 11593234]
- Luczak A, Hackett TA, Kajikawa Y, Laubach M. Multivariate receptive field mapping in marmoset auditory cortex. *J Neurosci Methods* 2004;136:77–85. [PubMed: 15126048]
- Malone BJ, Scott BH, Semple MN. Dynamic amplitude coding in the auditory cortex of awake rhesus macaques 10.1152/jn.01203.2006. *J Neurophysiol* 2007;98:1451–1474. [PubMed: 17615123]
- McAlpine D, Jiang D, Palmer AR. A neural code for low-frequency sound localization in mammals. *Nat Neurosci* 2001;4:396–401. [PubMed: 11276230]
- Merzenich MM, Brugge JF. Representation of the cochlear partition of the superior temporal plane of the macaque monkey. *Brain Res* 1973;50:275–96. [PubMed: 4196192]
- Middlebrooks JC, Xu L, Eddins AC, Green DM. Codes for sound-source location in nontopographic auditory cortex. *J Neurophysiol* 1998;80:863–81. [PubMed: 9705474]
- Miles R. Frequency dependence of synaptic transmission in nucleus of the solitary tract in vitro. *J Neurophysiol* 1986;55:1076–90. [PubMed: 3012009]
- Morel A, Rouiller E, de Ribaupierre Y, de Ribaupierre F. Tonotopic organization in the medial geniculate body (MGB) of lightly anesthetized cats. *Exp Brain Res* 1987;69:24–42. [PubMed: 3436391]
- Nagarajan SS, Cheung SW, Bedenbaugh P, Beitel RE, Schreiner CE, Merzenich MM. Representation of spectral and temporal envelope of twitter vocalizations in common marmoset primary auditory cortex. *J Neurophysiol* 2002;87:1723–37. [PubMed: 11929894]
- Nelken I, Versnel H. Responses to linear and logarithmic frequency-modulated sweeps in ferret primary auditory cortex. *Eur J Neurosci* 2000;12:549–62. [PubMed: 10712634]
- Nelken I, Chechik G, Mscis-Flogel TD, King AJ, Schnupp JW. Encoding stimulus information by spike numbers and mean response time in primary auditory cortex. *J Comput Neurosci* 2005;19:199–221. [PubMed: 16133819]
- Newman JD, Wollberg Z. Multiple coding of species-specific vocalizations in the auditory cortex of squirrel monkeys. *Brain Res* 1973;54:287–304. [PubMed: 4196750]
- Nicolelis MA, Ghazanfar AA, Stambaugh CR, Oliveira LM, Laubach M, Chapin JK, Nelson RJ, Kaas JH. Simultaneous encoding of tactile information by three primate cortical areas. *Nat Neurosci* 1998;1:621–30. [PubMed: 10196571]
- Orduna I, Mercado E 3rd, Gluck MA, Merzenich MM. Spectrotemporal sensitivities in rat auditory cortical neurons. *Hear Res* 2001;160:47–57. [PubMed: 11591490]
- Petersen RS, Panzeri S, Diamond ME. Population coding in somatosensory cortex. *Curr Opin Neurobiol* 2002;12:441–7. [PubMed: 12139993]
- Phillips DP, Kelly JB. Coding of tone-pulse amplitude by single neurons in auditory cortex of albino rats (*Rattus norvegicus*). *Hear Res* 1989a;37:269–79. [PubMed: 2708149]
- Phillips DP, Hall SE, Hollett JL. Repetition rate and signal level effects on neuronal responses to brief tone pulses in cat auditory cortex. *J Acoust Soc Am* 1989b;85:2537–49. [PubMed: 2745878]
- Rauschecker JP. Cortical processing of complex sounds. *Curr Opin Neurobiol* 1998;8:516–21. [PubMed: 9751652]
- Rauschecker JP, Tian B. Mechanisms and streams for processing of “what” and “where” in auditory cortex. *Proc Natl Acad Sci U S A* 2000;97:11800–6. [PubMed: 11050212]
- Rauschecker JP, Tian B. Processing of band-passed noise in the lateral auditory belt cortex of the rhesus monkey. *J Neurophysiol* 2004;91:2578–89. [PubMed: 15136602]

- Rauschecker JP, Tian B, Hauser M. Processing of complex sounds in the macaque nonprimary auditory cortex. *Science* 1995;268:111–4. [PubMed: 7701330]
- Rauschecker JP, Tian B, Pons T, Mishkin M. Serial and parallel processing in rhesus monkey auditory cortex. *J Comp Neurol* 1997;382:89–103. [PubMed: 9136813]
- Recanzone GH. Spatial processing in the auditory cortex of the macaque monkey. *Proc Natl Acad Sci U S A* 2000;97:11829–35. [PubMed: 11050216]
- Recanzone GH, Beckerman NS. Effects of intensity and location on sound location discrimination in macaque monkeys. *Hear Res* 2004;198:116–24. [PubMed: 15567608]
- Recanzone GH, Guard DC, Phan ML. Frequency and intensity response properties of single neurons in the auditory cortex of the behaving macaque monkey. *J Neurophysiol* 2000a;83:2315–31. [PubMed: 10758136]
- Recanzone GH, Guard DC, Phan ML, Su TK. Correlation between the activity of single auditory cortical neurons and sound-localization behavior in the macaque monkey. *J Neurophysiol* 2000b;83:2723–39. [PubMed: 10805672]
- Romanski LM, Bates JF, Goldman-Rakic PS. Auditory belt and parabelt projections to the prefrontal cortex in the rhesus monkey. *J Comp Neurol* 1999a;403:141–57. [PubMed: 9886040]
- Romanski LM, Averbach BB, Diltz M. Neural representation of vocalizations in the primate ventrolateral prefrontal cortex. *J Neurophysiol* 2005;93:734–47. [PubMed: 15371495]
- Romanski LM, Tian B, Fritz J, Mishkin M, Goldman-Rakic PS, Rauschecker JP. Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nat Neurosci* 1999b;2:1131–6. [PubMed: 10570492]
- Schreiner CE, Urbas JV. Representation of amplitude modulation in the auditory cortex of the cat. I. The anterior auditory field (AAF). *Hear Res* 1986;21:227–41. [PubMed: 3013823]
- Schreiner CE, Urbas JV. Representation of amplitude modulation in the auditory cortex of the cat. II. Comparison between cortical fields. *Hear Res* 1988;32:49–63. [PubMed: 3350774]
- Schreiner CE, Raggio MW. Neuronal responses in cat primary auditory cortex to electrical cochlear stimulation. II. Repetition rate coding. *J Neurophysiol* 1996;75:1283–300. [PubMed: 8867137]
- Schroeder CE, Lindsley RW, Specht C, Marcovici A, Smiley JF, Javitt DC. Somatosensory input to auditory association cortex in the macaque monkey. *J Neurophysiol* 2001;85:1322–7. [PubMed: 11248001]
- Schulze H, Langner G. Auditory cortical responses to amplitude modulations with spectra above frequency receptive fields: evidence for wide spectral integration. *J Comp Physiol [A]* 1999;185:493–508.
- Smiley JF, Hackett TA, Ulbert I, Karmas G, Lakatos P, Javitt DC, Schroeder CE. Multisensory convergence in auditory cortex, I. Cortical connections of the caudal superior temporal plane in macaque monkeys. *J Comp Neurol* 2007;502:894–923. [PubMed: 17447261]
- Stecker GC, Middlebrooks JC. Distributed coding of sound locations in the auditory cortex. *Biol Cybern* 2003a;89:341–9. [PubMed: 14669014]
- Stecker GC, Harrington IA, Middlebrooks JC. Location coding by opponent neural populations in the auditory cortex. *PLoS Biol* 2005;3:e78. [PubMed: 15736980]
- Stecker GC, Mickey BJ, Macpherson EA, Middlebrooks JC. Spatial sensitivity in field PAF of cat auditory cortex. *J Neurophysiol* 2003b;89:2889–903. [PubMed: 12611946]
- Steinschneider M, Reser DH, Fishman YI, Schroeder CE, Arezzo JC. Click train encoding in primary auditory cortex of the awake monkey: evidence for two mechanisms subserving pitch perception. *J Acoust Soc Am* 1998;104:2935–55. [PubMed: 9821339]
- Sugihara T, Diltz M, Averbach B, Romanski L. Integration of auditory and visual communication information in the primate ventrolateral prefrontal cortex. *J Neurosci* 2006;26:11138–11147. [PubMed: 17065454]
- Sutor B, Hablitz JJ. EPSPs in rat neocortical neurons in vitro. I. Electrophysiological evidence for two distinct EPSPs. *J Neurophysiol* 1989;61:607–20. [PubMed: 2709103]
- Tian B, Reser D, Durham A, Kustov A, Rauschecker JP. Functional specialization in rhesus monkey auditory cortex. *Science* 2001;292:290–3. [PubMed: 11303104]

- Wang X, Kadia SC. Differential representation of species-specific primate vocalizations in the auditory cortices of marmoset and cat. *J Neurophysiol* 2001;86:2616–20. [PubMed: 11698547]
- Wang X, Merzenich MM, Beitel R, Schreiner CE. Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: temporal and spectral characteristics. *J Neurophysiol* 1995;74:2685–706. [PubMed: 8747224]
- Whitfield IC, Evans EF. Responses of auditory cortical neurons to stimuli of changing frequency. *J Neurophysiol* 1965;28:655–72. [PubMed: 14347626]
- Wollberg Z, Newman JD. Auditory cortex of squirrel monkey: response patterns of single cells to species-specific vocalizations. *Science* 1972;175:212–4. [PubMed: 4621483]
- Woods TM, Lopez SE, Long JH, Rahman JE, Recanzone GH. Effects of stimulus azimuth and intensity on the single-neuron activity in the auditory cortex of the alert macaque monkey. *J Neurophysiol* 2006;96:3323–37. [PubMed: 16943318]

Fig. 1.

Tuning bandwidth for tones of MUs in this study, presented at 72 dB SPL. Horizontal gray lines span the responsive frequency range of each MU in ascending order of the upper edge frequency for low CF MUs and lower edge frequency of high CF MUs. Vertical dashed and solid lines indicate the frequency range of the FM₂₋₁₀ and FM₆₋₁₄ stimuli, respectively. Crosses and filled black squares indicate CF at threshold and center frequency at 72 dB SPL, respectively, for each MU, except for broadly tuned MUs that responded to tones over a 6 octaves range at the lowest intensity examined (from Kajikawa et al., 2005).

Fig. 2.

Response pattern of an MU with only phase-locked transients to FM₂₋₁₀. Data from MU number 14 in Fig. 1. (A) Raster plot, in which each row represents a single trial and every 10 rows were sorted by the repetition rate from 1 Hz (top) to 24 Hz (bottom). (B) PSTHs (bin-size: 8 ms) at repetition rates from 1 Hz (top) to 24 Hz (bottom). In both (A) and (B), the horizontal axis begins 50 ms before and ends 1600 ms after the onset of the first sweep. Light gray areas indicate sweep duration. (C) Entrainment pattern: the magnitude of phasic firing rate during each sweep (gray areas in A or B), normalized by the mean firing rate to the first sweep at all repetition rates, plotted by the order of sweeps in a train. Repetition rates are represented by gray scale intensity, from 1 Hz (lightest) to 24 Hz (darkest). (D) vsMTF (black

line with dots, left vertical axis) and rMTF (gray curve depicting mean and SD of 10 trials at each repetition rate, right axis). (E) PSTH similarity, evaluated for each repetition rate pair (pixels) by cross-correlation. Similarity is expressed by gray scale, between zero (black) and one (white). Diagonal pixels are auto-correlation coefficients (= 1) and the matrix is symmetrical around the diagonal. (F) The probability of classifying an observed response at a repetition rate (column) into different repetition rates (rows) by LDA. Log-transformed probability is expressed by gray scale, from zero (black) to one (white). Lighter pixel shading indicates higher probability of classifying the *column* repetition rate into the *row* repetition rate. (G) The probability of error classification (sum of off-diagonal probabilities within a column in F) at each repetition rate. Mean error rate (MER) is indicated by the *horizontal dashed line*. X-axis applies to panels E through G.

Fig. 3. MU response pattern to FM₆₋₁₄ exhibiting loss of the phase peak at high repetition rates. Data from MU number 17 in Fig. 1. Arrowheads in (D) label repetition rates where response synchronization was not significant. All other conventions were the same as Fig. 2.

Fig. 4. MU response pattern to FM₆₋₁₄ exhibiting loss of the phase peak at high repetition rates. Data from MU number 31 from Fig. 1. MER was 0.216. All other conventions were same as Fig. 2, except for the highest repetition rate examined was 20 Hz.

Fig. 5. Grand means of the normalized rMTF from MU responses to FM_{2-10} (*left*) and FM_{6-14} (*right*). Error bars indicate SD across MUs.

Fig. 6. vsMTF and response parameters derived by VS for FM₂₋₁₀ (left column) and FM₆₋₁₄ (right column). (A) Mean vsMTF and SDs across MUs. (B) Distribution of vsBMF. (C) Distribution of maxF.

Fig. 7.

(A) Curves showing mean classification error rate at each repetition rate. Error bars indicate SD across MUs in CM. (B) Distribution of BCF. In each panel, left and right side is for FM₂₋₁₀ and FM₆₋₁₄, respectively.

Fig. 8. (A) BCF is plotted as a function of vsBMF. (B) The difference, BCF - vsBMF, is plotted against vsBMF. In each panel, left and right side is for FM₂₋₁₀ and FM₆₋₁₄, respectively.

Fig. 9.

(A) Sonogram of marmoset twitter call A (tw A). Horizontal black and dashed lines demarcate the frequency ranges of FM₂₋₁₀ and FM₆₋₁₄, respectively. (B) PSTHs of the response to tw A recorded from 4 MUs (panels *a* – *d*). Horizontal axes match the bottom panel *d*. The ordinate of panel *b* applies to panels *a*–*c*. (C) Black lines show the VS spectrum for each of the 4 MUs (panels *a* – *d*). Gray lines show the envelope modulation spectrum of the sonogram shown in A. Dashed lines show the upper tail of 95 % confidence interval of the null structured (shuffled spike train) response spectrum. Both horizontal and vertical axes are same for all panels. The MUs depicted in panels *a*–*c* correspond to units depicted in Figs. 2, 3 and 4, respectively.

Fig. 10.

Top panels, sonogram of tw B (*upper*), and PSTH of response of one MU (*lower*). This MU is the same as in Fig. 9B (panel *b*) and Fig. 3. *Middle panels*, sonogram of half tw B (*upper*) and PSTH of same MU to half tw B (*lower*). Axes of *top* and *middle* figures are the same. *Bottom panels*, power spectra of the responses to tw B (*black line*) and half tw B (*gray line*). *Dotted lines* show upper and lower 95 % confidence intervals of spectra, estimated by bootstrap sampling of 15 trials.

Fig. 11.

(A) Waveforms of 3 twitter calls. (B, *a-c*) Raster plot (*top*) and PSTH (*bottom*) of 3 MUs. All graphs in A and B have identical time axes. (C, *a-c*) Bar graphs of classification error rates for 3 twitter calls by the 3 MUs in (B). Horizontal lines superimposed on bars indicate the VS spectral peak magnitudes of the response to each call. The error rate for one MU in B and C was zero for all calls.

Fig. 12.

Plot of the VS spectral peak at the tEMF of each call against classification error rate for MUs with significant entrainment at the tEMF. tw A, filled circle; tw B, open circle; tw C, square.

Table 1

Spearman's rank correlation coefficients for two sets of comparisons: *VS*, correlation between VS spectral peak magnitudes of the twitter call response and VS of the FM sweep train response at the tEMF of each twitter call; *Error rate*, correlation between classification error rate of the twitter call and repetition rate of the FM sweep train at the tEMF of each all.

	tw (tEMF Hz)	VS	Error rate
FM ₂₋₁₀ (<i>n</i> = 20)	A (6.5)	-0.0006	0.4983 [*]
	B (8.5)	0.4436 [*]	0.6091 ^{**}
	C (7.0)	0.4150 [*]	0.5829 ^{**}
FM ₆₋₁₄ (<i>n</i> = 35)	A (6.5)	0.3227 [*]	0.5885 ^{**}
	B (8.5)	0.1451	0.3488 [*]
	C (7.0)	0.3515 [*]	0.4704 ^{**}

^{*}
 $P < 0.05$

^{**}
 $P < 0.005$.