UCSF
UC San Francisco Previously Published Works

Title

Modulation-Frequency-Specific Adaptation in Awake Auditory Cortex

Permalink

Journal

ISSN

Authors

Malone, Brian J
Beitel, Ralph E
Vollmer, Maike
et al.

Publication Date

DOI

Peer reviewed

Systems/Circuits

Modulation-Frequency-Specific Adaptation in Awake Auditory Cortex

Brian J. Malone,[1,3]* **Ralph E. Beitel,**[1,3]* **Maike Vollmer,**[4] **Marc A. Heiser,**[5] **and Christoph E. Schreiner**[1,2,3]

[1]Coleman Memorial Laboratory, [2]Center for Integrative Neuroscience, and [3]Department of Otolaryngology-Head and Neck Surgery, University of California, San Francisco, California 94143, [4]Comprehensive Hearing Center, Department of Otolaryngology-Head and Neck Surgery, University Hospital Wuerzburg, 97080 Wuerzburg, Germany and [5]Department of Psychiatry, Child and Adolescent Division, University of California Semel Institute for Neuroscience and Behavior, Los Angeles, California 90024

Amplitude modulations are fundamental features of natural signals, including human speech and nonhuman primate vocalizations. Because natural signals frequently occur in the context of other competing signals, we used a forward-masking paradigm to investigate how the modulation context of a prior signal affects cortical responses to subsequent modulated sounds. Psychophysical "modulation masking," in which the presentation of a modulated "masker" signal elevates the threshold for detecting the modulation of a subsequent stimulus, has been interpreted as evidence of a central modulation filterbank and modeled accordingly. Whether cortical modulation tuning is compatible with such models remains unknown. By recording responses to pairs of sinusoidally amplitude modulated (SAM) tones in the auditory cortex of awake squirrel monkeys, we show that the prior presentation of the SAM masker elicited persistent and tuned suppression of the firing rate to subsequent SAM signals. Population averages of these effects are compatible with adaptation in broadly tuned modulation channels. In contrast, modulation context had little effect on the synchrony of the cortical representation of the second SAM stimuli and the tuning of such effects did not match that observed for firing rate. Our results suggest that, although the temporal representation of modulated signals is more robust to changes in stimulus context than representations based on average firing rate, this representation is not fully exploited and psychophysical modulation masking more closely mirrors physiological rate suppression and that rate tuning for a given stimulus feature in a given neuron's signal pathway appears sufficient to engender context-sensitive cortical adaptation.

Key words: auditory; context; monkey; primate; suppression; unanesthetized

Introduction

Amplitude modulation (AM) is a ubiquitous feature of complex sounds, particularly communication sounds. AM is crucial for speech intelligibility—vocoded speech demonstrates the sufficiency of envelope cues for intelligibility (Shannon et al., 1995), modulation sensitivity is one of the best predictors of speech comprehension (Cazals et al., 1994, Fu, 2002; Won et al., 2011), and the AM rate can drive the segmentation of complex auditory scenes (Grimault et al., 2002; Dolležal et al., 2012). Measuring modulation sensitivity is central to psychoacoustics, in which temporal resolution is often defined by the temporal modulation transfer function (tMTF), which characterizes the minimum de-

tectable modulation depth as a function of the modulation frequency (Viemeister, 1979).

Critically, the tMTF depends on the context in which stimuli are presented. For example, the threshold for detecting sinusoidal amplitude modulation (SAM) is higher when a modulated "probe" signal is preceded by a "masker" signal modulated at a similar modulation frequency (Wojtczak and Viemeister, 2005). Prior observations of "modulation masking" for both frequency (Kay and Matthews, 1971, 1972) and amplitude (Bacon and Grantham, 1989; Houtgast, 1989; Dau et al., 1997; Kohlrausch et al., 2000) modulations suggested that there are channels, or filters, for detecting modulation frequency analogous to those for detecting spectral frequency. Computational models of the auditory system have incorporated a modulation filterbank (Dau et al., 1997; McDermott and Simoncelli, 2011) to "reflect the auditory system's high sensitivity to fluctuating sounds and to account for amplitude modulation (AM) detection and masking data" (Jepsen et al., 2008).

How such a filterbank is implemented physiologically remains unclear, but the band-pass parameters of the modulation filterbank of Dau et al. (1997) were inspired by physiological studies of the inferior colliculus (IC) (Langner and Schreiner, 1988; Langner, 1992). Importantly, MTFs measured in the auditory nerve are essentially low pass and are substantially flat within the pass

band (Joris and Yin, 1992). Prominent rate tuning does not begin to emerge until the IC (Creutzfeldt et al., 1980; Schreiner and Urbas, 1986, 1988; Langner and Schreiner, 1988; Langner, 1992; Krishna and Semple, 2000). Therefore, modulation frequency tuning appears to be an emergent phenomenon in the ascending auditory system.

A recent study of modulation masking found that relatively few IC neurons exhibited responses consistent with modulation masking (Wojtczak et al., 2011), leaving open the possibility that modulation masking has stronger physiological parallels with more central auditory structures such as the cortex. Numerous physiological studies have begun to clarify the cortical representations of modulated signals (for review, see Joris et al., 2004; Malone and Schreiner, 2010), including substantial work in primate models (Bieser and Müller-Preuss, 1996; Lu et al., 2001; Liang et al., 2002; Bartlett and Wang, 2005; Malone et al., 2007, 2010, 2013, 2014; Johnson et al., 2012). Here, we searched for cortical evidence of modulation masking in a forward-masking paradigm with SAM applied to tonal carriers in awake squirrel monkeys. We evaluate evidence for cortical modulation filter-banks and discuss the implications of modulation-frequency-specific adaptation for perception.

Materials and Methods

Surgical preparation. All procedures related to the maintenance and use of animals in this study were approved by the Institutional Animal Care and Use Committee of the University of California–San Francisco (UCSF) and followed guidelines of the National Institutes of Health for the care and use of laboratory animals. The methodological details for these experiments have been described previously (Malone et al., 2013), but some are repeated here for the readers' convenience.

Two adult female squirrel monkeys (*Saimiri sciureus*) were trained to sit quietly in a restraint chair. Animals were then implanted with head posts to allow for head fixation during physiological recording. During all surgical procedures, anesthesia was induced with ketamine (25 mg/kg, i.m.) and midazolam (0.1 mg/kg) and the animals were maintained in a steady plane of anesthesia using isoflurane gas (0.5–5%). Implants were secured to the skull using bone screws and dental acrylic. After animals were trained to sit in the primate chair with their head fixed to a frame, they underwent a second surgery to implant a recording chamber over auditory cortex. The temporal muscle was resected, the cranium overlying auditory cortex was exposed, and a 10-mm-diameter ring was secured using bone screws and dental acrylic. Perioperative pain management included local application of bupivacaine, as well as buprenorphine (0.01–0.03 mg/kg and meloxicam (0.3 mg/kg) as needed, and in consultation with veterinary staff in the UCSF Laboratory Animal Resource Center.

Sterile procedures were used to expose and record from auditory cortex. A 2–3 mm burr hole was drilled using either a dental drill mounted on a micromanipulator under magnification with a surgical microscope or using a hand drill. A small incision was made in the dura using microsurgical instruments after application of a drop of 1% lidocaine. After several recording sessions in a burr hole, another burr hole was drilled and the recording process was repeated. Burr holes were sometimes enlarged or connected by removing bone with fine surgical instruments after application of lidocaine as needed to expose additional areas of auditory cortex. After each recording session, the chamber was filled with antibiotic ointment and sealed with a metal cap.

Electrophysiology. All recordings were made in a soundproof chamber (Industrial Acoustics). During each recording session, the animal was seated comfortably in a custom-built primate chair with its head fixed to a frame while stimuli were presented. Data were obtained using 16-channel linear electrodes (177 μm^2 contact size, 100 or 150 μm spacing) from NeuroNexus Technologies. An electrode was advanced into cortex using a microdrive (David Kopf Instruments) to the depth at which most channels were active (tip depth of ~1–2 mm from the depth of first spontaneous activity identified audiovisually). Penetrations were made

approximately perpendicular to the surface of the exposed cortex, but it was not always possible to achieve electrode orientations orthogonal to the cortical surface in some recording locations. Recording sessions typically lasted ~2.5 h.

Electrical signals from the brain were amplified using a 16-channel preamplifier (RA16 Medusa; Tucker-Davis Technologies), band-pass filtered (600–7000 Hz) and recorded using an RX-5 amplifier and Brainware software (Tucker-Davis Technologies) on a personal computer. Brainware was used for online estimation of neural responsiveness and tuning and raw waveforms were sampled (25 kHz).

In the squirrel monkey, the core auditory fields (primary auditory cortex: AI; field R: R) are located on the surfaces of the temporal gyrus and in the supratemporal plane of the lateral sulcus. The location of our recordings within auditory cortex was determined physiologically by the characteristics of core auditory neurons, including vigorous pure tone responses, short response latencies, and a tonotopic gradient in the rostrocaudal dimension (Cheung et al., 2001; Scott et al., 2011; Malone et al., 2013).

Stimulus delivery. All sounds in this study were presented using a field speaker (Sony SS-MB150H) placed directly in front of the animal. Distance from the front of the speaker to the interaural line was 40 cm. The sound delivery system was calibrated using a sound-level meter and Sig-Cal software (Tucker-Davis Technologies). Levels were measured using a Brüel and Kjær model 2209 meter, an A-weighted decibel filter, and a model 4192 microphone. Levels in the initial hemisphere varied over a range of 62–72 dB, with an average of 66.8 dB. Levels in the remaining two hemispheres were more tightly constrained, from 64 to 66 dB, with averages of 64.8 and 65.4, respectively, and all sound levels within the same recording session were within 1 dB of each other.

Sinusoidal amplitude modulation. The sinusoidal amplitude modulation (SAM) signals described in this report consisted of a sinusoidal carrier tone (f_c) modulated sinusoidally by a second tone at a lower frequency (f_m) such that $s(t) = A[1 + m \cdot \sin(2\pi f_m t + \Phi)]\sin(2\pi f_c t)$, where $s(t)$ is the signal and t refers to time. The phase term, Φ, was equal to $-\pi/2$, so that each modulation cycle begins and ends at the minimum amplitude within the cycle. For all stimuli, the modulation depth (m) was set to 100%. SAM signals with pure tone carriers are uniquely defined by only four parameters: carrier frequency (f_c), modulation frequency (f_m), carrier level (A), and modulation depth (m). Online estimates of the best frequency (BF) were used to select the carrier frequency for SAM (see below).

We presented masker signals immediately followed by probe signals. This forward-masking paradigm obviates potential confounds in simultaneous masking, such as local temporal cues (e.g., "dip listening"), beating due to peripheral nonlinearities, and modulation distortion products (Wojtczak and Viemeister, 2005). Figure 1 depicts the organization of the experiment. Data were typically collected in a series of trials lasting 2 s. Masker signals of 1000 ms duration were modulated at 1 of 4 masker modulation frequencies (4, 10, 32, or 96 Hz). When the masker signal was an unmodulated carrier tone, we refer to it the "0 Hz masker" for notational convenience. Maskers were immediately followed by 1000 ms probe signals modulated at one of the frequencies comprising the MTF (4, 6, 8, 10, 16, 24, 32, 64, 96, 128, 192, 256, 384, and 512 Hz). In some experiments, slightly different frequencies were used, but only values similar to those in the foregoing list were included for analysis (e.g., 250 vs 256 Hz). Only modulation frequencies that matched exactly were compared directly across different stimulus contexts (see below). Trials were separated by interstimulus intervals of 500 or 600 ms.

Sets of trials for a given masker (e.g., 4 Hz) were presented in blocks of "runs." SAM stimuli for a given masker were presented in pseudorandom order until 20 trials had been presented at each modulation frequency. Each run lasted ~13 min and, in some cases, runs involving changes in the spectrum of the carrier (Malone et al., 2013) were interspersed with runs in which the masker modulation frequency was varied. As a result, the entire duration of the recording session could be quite long (>1.5 h). For this reason, we report multiunit responses because tracking individual spike waveforms over such long intervals proved impractical. We defined multiunit activity at a given recording site as the time points where the filtered voltage waveform exceeded 3.5 SDs of its

amplitude distribution. By collecting the data in this way, we maximized our ability to compare responses across the five maskers (0, 4, 10, 32, and 96 Hz) at each site. Because all possible comparisons (e.g., 4 Hz vs 96 Hz) were generally not available for each site, the total number of sites for each comparison are not necessarily equal.

Carrier frequencies for both the maskers and probes were chosen to reflect the modal BF among the active channels in each electrode penetration based on estimates of the BF obtained online. Use of the modal BF for each penetration was necessary because all penetrations were not strictly perpendicular to the cortical surface and the BF could vary across electrodes at different depths in the array. In most cases, however, the BFs were fairly consistent across all electrodes in each penetration. The carrier frequencies used in the experiments varied from 0.5 to 20 kHz. When tallied as kHz (count), they were as follows: 0.5 (2), 1 (1), 2 (2), 3 (3), 4 (9), 6 (2), 10 (2), 12 (2), 16 (1), and 20 (1). Frequency tuning for each channel was estimated offline from responses to tone pips drawn from a standardized list of frequency and level combinations spanning multiple octaves and an amplitude range of 0–70 dB in 10 dB steps (Malone et al., 2013).

Modulation analysis. All data analysis was performed using MATLAB (MathWorks). The MTFs were analyzed with respect to average firing rate and vector strength. We refer to the MTFs describing the changes in rate and VS as rMTFs and vsMTFs, respectively. The rMTFs were calculated by averaging the spike counts obtained during each probe presentation across all trials ($n = 20$). Significant differences in firing rate across different modulation frequencies were determined by comparing the distributions of spike rates across trials using a Wilcoxon rank-sum test. A MTF was considered to exhibit modulation rate tuning if the difference in average firing rate between the maximum and minimum values was significant ($p < 0.001$).

Vector strength (VS) (Goldberg and Brown, 1969) was used to measure the degree to which the neural response was concentrated at a particular phase of the modulation cycle, such that $VS = (1/n) \cdot \Sigma(\cos(2\pi \cdot f_m \cdot t_i)^2 + \sin(2\pi \cdot f_m \cdot t_i)^2)^{0.5}$, where t_i is the time of occurrence of the ith spike, n is the total number of spikes, and f_m is the modulation frequency. A neuron was considered to be synchronized to the modulation envelope if the Rayleigh statistic, $2 \cdot VS^2 \cdot n$, exceeded 13.816 (corresponding to $p < 0.001$; Mardia and Jupp, 2000). If all spikes occur at the same modulation phase, then $VS = 1$. If all spikes are evenly distributed in the modulation cycle, $VS = 0$. To stabilize the estimate, we performed this procedure using half the trials 10 times and stored the average value. To limit the effects of multiple onsets, the initial 100 ms of data from each trial were eliminated from the VS calculations. A vsMTF was considered to exhibit significant synchronization if the VS was significant by the Rayleigh test for at least one tested modulation frequency. Electrode channels that failed to exhibit either significant rate tuning or synchronization for any tested SAM stimuli were considered unresponsive and eliminated from further analysis.

Generation of composite difference functions. Our goal was to characterize masker effects on complete rMTFs and vsMTFs. To visualize general trends in the data at the population level, we constructed composite difference functions (CDFs). For each penetration, we attempted to present as many different masking conditions as feasible given the recording conditions. We identified all possible pairs of complete MTFs obtained in the context of different maskers (e.g., 4 vs 32 Hz). For comparisons that included the 0 Hz masker, we subtracted the MTF for the 0 Hz (i.e., unmodulated) masker from the modulated masker. For pairs that included maskers at two different modulated frequencies, we subtracted the MTF associated with the lower-frequency maskers from the MTF associated with the higher-frequency masker. The subtraction resulted in what we term difference functions. For rMTFs, we expressed the differ-
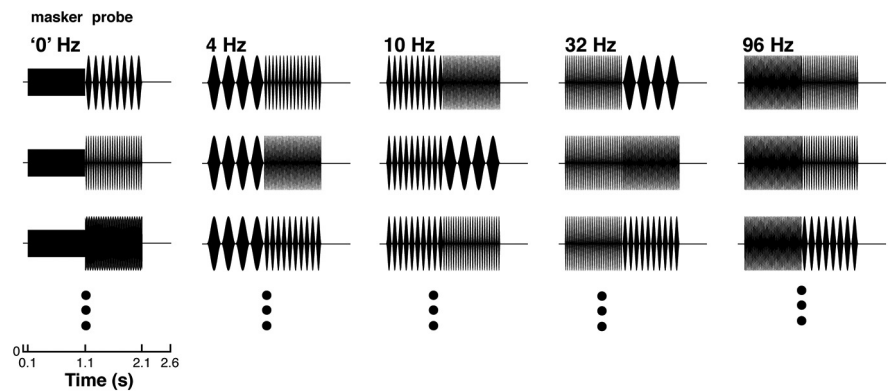

Figure 1. General organization of the experiment. Each column contains a diagram showing the stimuli for a 3 stimulus trial for 1 of the 5 masking conditions. The 0 Hz masker is unmodulated (leftmost column) and the remaining modulated maskers were fully (100%) modulated at the frequencies indicated above the icons. The maskers were immediately followed by the SAM signals. Different maskers were presented in blocks. Each block contained 20 repetitions of each modulation frequency comprising the MTF presented in pseudorandom order preceded by the masker for that block.

ence relative to the highest firing rate obtained on either rMTF. This normalization was intended to prevent sites with the highest firing rates from dominating the averages. Difference functions based on VS were not normalized because the VS metric is bounded from 0 to 1. Difference functions based on a particular comparison (e.g., a 4 Hz masker versus a 32 Hz masker, or 32 Hz–4 Hz) were identified across all recording sites (including different electrode channels within the same penetration), and averaged together to form the CDFs.

Given this arrangement, the same MTF recurs in multiple analyses. For example, in a penetration where the 4, 32, and 96 Hz maskers were presented, the MTF associated with the 32 Hz masker would be subtracted from the MTF associated with the 96 Hz masker and the resulting difference function would then be averaged in the 96 Hz–32 Hz CDF; the same MTF will also have the MTF associated with the 4 Hz masker subtracted from it to form a difference function contributing to the 32 Hz–4 Hz CDF. This subtractive analysis is necessary to preserve the differential effects of modulation frequency masking because cortical MTFs in the awake squirrel monkey are heterogeneous (Malone et al., 2013) and averaging MTFs obtained in different contexts across different neural clusters obscures the effects of interest. Because the order of the presentation of experimental blocks (Fig. 1) varied from penetration to penetration, the number of difference functions contributing to each CDF varies, from as few as 43 recording sites for the 4 Hz–0 Hz CDF to as many as 249 sites for the 10 Hz–4 Hz CDF. The average site count was 152.2 sites per CDF.

Statistical verification of contextual modulation effects. Verification of significant changes in the MTFs due to differences in the masker SAM were quantified on a site-by-site basis using the procedure described in Malone et al. (2013). We treated each MTF as a vector with entries corresponding to each tested modulation frequency. Differences between MTFs were quantified by a similarity index (SI), defined as the (L2) vector norm of the two MTFs divided by the sum of their respective vector norms: $1 - \|\text{mtf1} - \text{mtf2}\|/(\|\text{mtf1}\| + \|\text{mtf2}\|)$. To assign significance to this value, we created a set ($n = 1000$) of bootstrapped estimates for the SI by randomly assigning the trials from the two MTFs being compared to create two "blended" MTFs. Significance was assigned by counting the number of cases in which the SIs of "blended" MTFs exceeded the actual SI and then dividing by the number of iterations (i.e., if none did, $p < 0.001$). The logic of this test is that, if the responses on each trial are drawn from the same underlying distribution (i.e., responses to the SAM signals are the same for both maskers, subject to trial to trial variability), then it is unlikely that the actual set of measured trials will be maximally dissimilar. Therefore, a random reshuffling of the trials across carrier types should produce SI values lower than the actual SI value. If not, then the maskers likely did have an effect and the two MTFs sample data from distinct distributions.
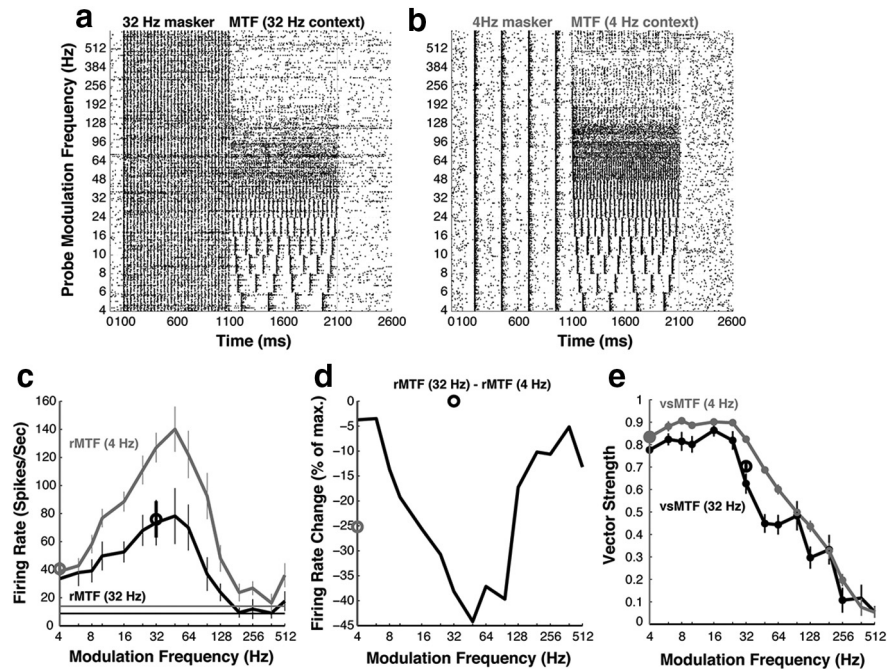
Figure 2. Example of the effects of modulation frequency context for a cluster of neurons at a single recording site. Each tick on the raster plots in **a** and **b** indicates the occurrence of a spike. Responses to each stimulus are stacked by trials. rMTFs based on averaging the firing rates for the test SAM stimuli are shown in **c**. The rMTF obtained in the context of the 32 Hz masker is shown in black and the rMTF obtained in the context of the 4 Hz masker is shown in gray. Vertical lines indicate ±2 SEM across repeated trials ($n = 20$). The larger circles indicate the average firing rates elicited by the masker SAM stimuli (averages were computed across all trials because every trial included the masker SAM). The curve in **d**, the difference function, indicates the difference in firing rates obtained by subtracting the gray curve in **c** from the black curve in **c** and dividing by the maximum firing rate observed on either rMTF. The circles indicate the difference in firing rates elicited by the maskers, normalized by the highest rate observed for either rMTF. **e**, vsMTFs computed based on the raster data shown in **a** (black curve) and **b** (gray curve). Vertical lines indicate ±1 SD for the VS values when the vsMTF was iteratively ($n = 50$) computed based on random draws of 10 trials. The plotted curves represent the mean VS across iterations. Filled circles on each curve indicate significant ($p < 0.001$) synchronization by the Rayleigh criterion (see Materials and Methods). The larger circles indicate the mean VS for the maskers across iterations.

Analysis of the CDFs fell into three main categories. First, we assessed the total suppression magnitude induced by the masker by directly comparing the distribution of values underlying two different CDFs that share a baseline (e.g., 32 Hz–4 Hz versus 96 Hz–4 Hz, where 4 Hz is the common baseline) with Wilcoxon rank-sum tests. This analysis is indifferent to how contextual modulatory effects are distributed across modulation frequency. For example, two CDFs might both be essentially flat, but offset relative to one another because one of the maskers more effectively suppresses responses to the probe regardless of the relationship between the masker and probe modulation frequencies.

The second analysis, in contrast, addresses whether any masker-induced response changes exhibit tuning across modulation frequency. We use the variance of the CDF as an indirect index of tuning because the variance of a flat CDF would be zero. CDF variance will be higher if the contextual influence of the maskers interacts with modulation tuning. Note that this analysis is indifferent to the "DC offset" of the CDF and complements the analysis of total suppression magnitude. We used permutation tests to compare the variance of each CDF with the variance expected for the empirical distribution of difference function values. To do so, we randomly reshuffle the values in each difference function before averaging them to generate a simulated CDF. We assign significance by counting how often the variance of simulated CDFs exceeded the actual CDF variance and dividing by the number of iterations ($n = 10,000$).

The third analysis addresses the locus of context-induced MTF tuning changes. For this analysis, we computed the suppressive center of mass (COM) for each CDF rather than relying on a single value (e.g., the CDF minimum). We define the suppressive COM as the product of the negative values of the CDF and their values in modulation frequency (ex-

pressed as base 2 logarithms, such that 16 Hz = 4), divided by the sum of those negative values. To verify that the COM differed between a pair of CDFs, we compared the difference in the actual COM values against simulated COM differences based on random mixtures of the data comprising the CDFs using the same logic explained above. When comparing two modulated maskers, positive CDF values for a given modulation frequency range could also reflect suppression; because the difference functions that were averaged to produce the CDF are based on subtracting one MTF from another. Suppression of the subtrahend relative to the minuend results in positive values (see Fig. 5f). Empirically, however, one masker tended to dominate the contextual effects we observed and the foregoing test appeared adequate to substantiate shifts in the dominant modulation frequency locus of suppression.

When analyzing population distributions of continuous variables, we compared median values via nonparametric Wilcoxon rank-sum tests unless otherwise stated. Correlations were quantified in terms of the Pearson product-moment coefficient.

Results

Summary of the data sample

The data in this report are derived from clusters of neurons recorded on 285 distinct channels during 25 penetrations using linear 16 channel probes. Penetrations were made into core auditory cortex located in three hemispheres of two alert adult squirrel monkeys. The data described herein substantially overlap with the data reported in Malone et al. (2013). We included all data when we were able to obtain complete MTFs for at least two masking conditions. We also required that each MTF exhibit either significant rate tuning or significant synchronization to at least one modulation frequency (see Materials and Methods). MTFs that did not meet this criterion were excluded from further analysis. The median number of masking conditions obtained per channel was 4 (mean of 3.74), which allowed us to generate difference functions for 1522 pairs of MTFs.

Examples of modulation-frequency-specific firing rate adaptation

Figure 2 contains examples of spike rasters for MTFs recorded in the context of maskers at 32 Hz (Fig. 2a) and 4 Hz (Fig. 2b). The 32 Hz masker elicited robust and highly synchronized responses from this site. The 4 Hz masker also elicited precisely synchronized responses, although at a somewhat lower firing rate. Values for the average firing rates elicited by the maskers are indicated by the large circles added to the rMTFs in Figure 2c. Presentation of the 32 Hz masker (black curve) significantly curtailed the responses to the probe SAM signals over a broad range of modulation frequencies relative to the 4 Hz masker (gray curve). We illustrate this effect more concisely in Figure 2d, which shows the difference between the rMTFs expressed as a percentage of the maximum firing rate obtained in either condition (i.e., for 48 Hz after the 4 Hz masker). It is critical to note that, within the temporal interval from 1.1 to 2.1 s, the stimuli in Fig. 2, a and b, are

strictly identical—the only difference is the prior stimulus context. If the maskers did not differentially affect the responses of the site, then the curve in Fig. 2d would be flat. As the graph makes clear, however, the greatest suppression occurs ~48 Hz, which is both near the peak of the rMTF for the 4 Hz masker and adjacent to the 32 Hz masker (Scholes et al., 2011).

Figure 3 shows the results obtained when responses subsequent to 96 Hz and 10 Hz maskers are compared. Comparison of the rasterplots in Figure 3, a and b, demonstrates that responses to the 96 Hz masker were weaker and less synchronized than responses to the 10 Hz masker. These observations can be confirmed by examination of the rMTFs (Fig. 3c) and vsMTFs (Fig. 3e). Nevertheless, exposure to the 96 Hz masker clearly reduces responses to probes near 96 Hz, a result evident in both the rasterplot (Fig. 3a) and in the sharp trough in the difference function shown in Figure 3d. In addition, exposure to the 10 Hz masker appears to reduce responses to probe SAM at modulation frequencies <48 Hz, suggesting that complementary, modulation-specific adaptation effects produce the striking "crossover" in the rMTFs (Fig. 3a). Unlike the example depicted in Figure 2, the trough in the difference function does not coincide with the best modulation frequency of the recording site (i.e., 512 Hz).

To determine the significance of modulation frequency context on a site by site basis, we used Monte Carlo techniques to determine whether the difference in the rMTFs obtained with different maskers exceeded the range of differences obtained for random mixtures of trials across the adapting conditions (see Materials and Methods). We quantified the difference between each pair of rMTFs with an SI and compared the actual SI with the distribution of simulated SIs. By this test, 56% (833/1480) of the comparisons across different maskers (including unmodulated tones) were significant (p < 0.001). If we normalize the curves by their sums before this analysis, then 21% (305/1480) of the comparisons remain significant (p < 0.001), indicating that modulation frequency context affected the shape as well as the scale of the rMTFs in a significant minority of cases.

Modulated maskers elicit significant tuned suppression of cortical responses
To assess the effects of modulation frequency context more generally, we constructed sets of CDFs by averaging the individual difference functions like those shown in Figures 2d and 3d across all clusters in the population. In the simplest cases, we compared maskers at a given modulation frequency (4, 10, 32, or 96 Hz) against maskers that consisted of unmodulated tones (0 Hz). Results of this analysis are depicted in Figure 4. By convention, we subtracted the rMTFs obtained after exposure to the unmodulated maskers from rMTFs obtained after exposure to the modulated maskers. Therefore, points <0 (indicated by a horizontal line in all panels) reflect suppression of the response attributable to the modulation. If the presentation of the maskers had either no effect or equivalent effects, then all curves would lie along the horizontal line.
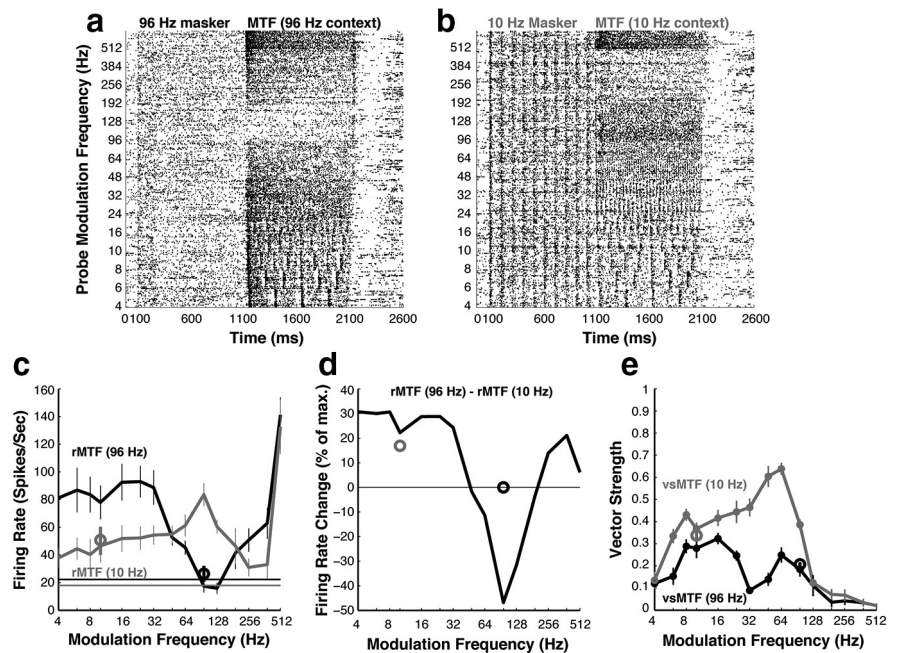

Figure 3. Figure conventions for the raster plots in *a* and *b* and the curves plotted in *c*, *d*, and *e* are the same as those used in Figure 2.

There are two main features of the CDFs in Figure 4. First, increases in the modulation frequency of the maskers produce greater suppression such that the CDFs shift down from left to right (Fig. 4a–d). We quantified this effect by computing the median value of all points comprising the CDFs, resulting in values of −1.1%, −2.7%, −5.2%, and −5.5% for the 4, 10, 32, and 96 Hz maskers, respectively. These medians were significantly different from 0 for maskers of 10 Hz and higher (p < 10^{-8}; Wilcoxon signed rank).

The second salient feature of the CDFs in Figure 4 is the fact that modulation frequency masking appears to be tuned. We determined whether the CDFs exhibited tuning by comparing the variance of the actual CDFs against simulated CDFs based on averages of randomly reshuffled difference functions (see Materials and Methods). If the variance of the actual CDFs did not exceed that of the simulated CDFs, we concluded that the apparent tuning of the CDF could have occurred by chance given the distribution of values comprising the difference functions. By this test, all the CDFs in Figure 4 exhibited significant tuning for modulation frequency (p < 0.0001).

We observed similar patterns when comparing results for two different modulated maskers, as shown in Figure 5. By convention, we subtracted the rMTF obtained after exposure to the lower-modulation-frequency masker from the rMTF obtained after exposure to the higher-modulation-frequency masker. All points on the curves in Figure 5, a and b, lie below zero, indicating that 10 Hz and 32 Hz suppressed the subsequent responses relative to the 4 Hz masker. The 96 Hz masker produced the greatest suppression, but the effect was more tightly constrained to higher modulation frequencies.

In general, the higher-modulation-frequency masker resulted in greater suppression. For example, the median value for the 10 Hz–4 Hz comparison was −3.5% compared with −5.3 for the 32 Hz–4 Hz comparison and −4.7% for the 96 Hz–4 Hz comparison. The exception was the 96 Hz–32 Hz comparison (Fig. 5d), for which the median value was 0, because the curve is effectively
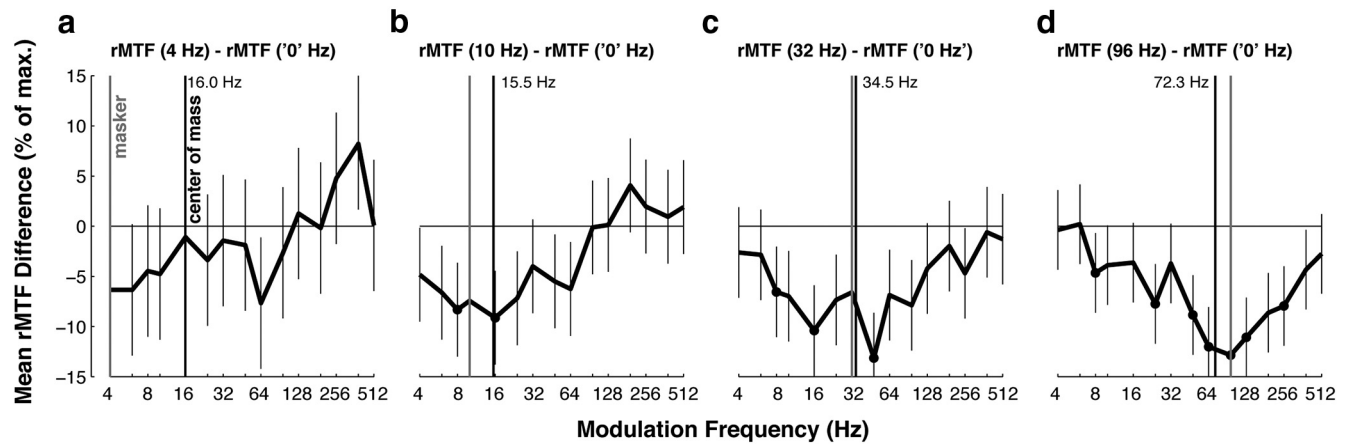
Figure 4. CDFs based on comparisons between modulated and unmodulated maskers. The identities of the rMTFs and the order of subtraction used to produce the depicted CDFs are indicated by the titles appearing above panels *a–d*. Each CDF represents the average of all difference functions (e.g., Figs. 2*d*,3d) for that particular comparison available in the data sample. Vertical gray lines indicate the modulation frequency of the modulated masker. The thick vertical black line indicates the suppressive COM for the CDF (see Materials and Methods). COM values are indicated adjacent to the line. The thin vertical black lines on each CDF indicate ±2 SEM. Filled circles on each curve indicate significant differences from a median of zero ($p < 0.001$; Wilcoxon signed rank).
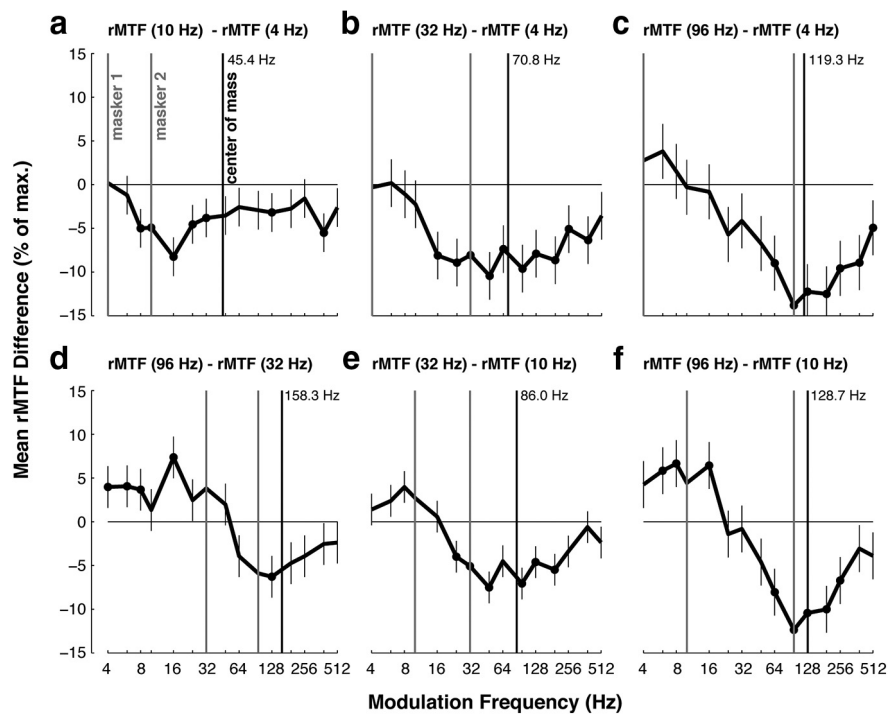

Figure 5. The identities of the rMTFs and the order of subtraction used to produce the depicted CDFs are indicated by the tiles appearing above panels *a–f*.

balanced, being positive below ∼48 Hz and negative above it. In this case, the 32 Hz masker suppresses responses at modulation frequencies <48 Hz more effectively than the 96 Hz masker, whereas the 96 Hz masker suppresses responses >48 Hz more effectively than the 32 Hz masker.

To demonstrate that total suppression was greater for higher modulation frequencies, we compared the distributions of differences (all points on the rMTFs from all sites) using the unmodulated control ("0" Hz) as a baseline. For example, we compared the 4 Hz–0 Hz distribution against the "96 Hz–0 Hz" distribution. For the comparisons involving two modulated maskers, we tested comparisons against a common baseline masker: 32 Hz–10 Hz versus 96 Hz–10 Hz, where 10 Hz serves as the baseline response. Examination of Table 1 shows that, when modu-

maskers are compared against a common baseline, the masker with a higher modulation frequency produced more robust suppression of the subsequent response. The exceptions are the cases where 96 Hz and 32 Hz maskers are compared against a common baseline, which were never significant, suggesting that increasing the modulation frequency beyond 32 Hz does not increase the magnitude of suppression. Finally, we verified that suppression was tuned for comparisons between two modulated maskers using the variance-based test described above. All of the CDFs in Figure 5 also exhibited significant tuning ($p < 0.0001$).

Locus of maximal suppression shifted to higher modulation frequencies for higher-modulation-frequency maskers
Having verified that the tuning evident in the CDFs is genuine, we tested whether the locus of greatest suppression occurred near the modulation frequency of the masker. To do so, we calculated the COM of the suppressed (i.e., negative) values of each CDF (see Materials and Methods). The suppressed COMs for each curve are indicated by black vertical lines in Figures 4 and 5. By analogy to spectral forward masking (Scholes et al., 2011; Zhou and Wang, 2014), we would expect that the locus of suppression would occur near the masker modulation frequency. For example, we would expect that the center of mass for the 96 Hz–10 Hz CDF to be shifted to the right of that for the 32 Hz–10 Hz CDF (responses after the 10 Hz masker serve as a baseline for the estimate).

To provide statistical verification of such tuning shifts, we computed the actual difference in the centers of mass for each pair of CDFs that include a common baseline against differences in the centers of mass based on simulated CDFs comprised by mixtures of the difference functions that were averaged to produce the actual CDFs (see Materials and Methods). Results of

Table 1. Results of analyses comparing the magnitude of response suppression across different CDFs

CDF$_1$	CDF$_2$	CDF$_1$ median (%)	CDF$_2$ median (%)	p-value
4 vs 0	10 vs 0	−1.12	−2.68	0.0321
4 vs 0	32 vs 0	−1.12	−5.25	$<10^{-5}$
4 vs 0	96 vs 0	−1.12	−5.50	$<10^{-6}$
10 vs 0	32 vs 0	−2.68	−5.25	= 0.0051
10 vs 0	96 vs 0	−2.68	−5.50	= 0.0013
32 vs 0	96 vs 0	−5.25	−5.50	= 0.7372
10 vs 4	32 vs 4	−3.52	−5.30	$<10^{-5}$
10 vs 4	96 vs 4	−3.52	−4.72	= 0.0024
32 vs 4	96 vs 4	−5.30	−4.72	= 0.2846
32 vs 10	96 vs 10	−1.92	−1.44	= 0.2569
32 vs 10	32 vs 4	−1.92	−5.30	$<10^{-8}$
96 vs 4	96 vs 10	−4.72	−1.44	<0.0003
96 vs 32	96 vs 10	0	−1.44	$<10^{-22}$
96 vs 32	96 vs 4	0	−4.72	$<10^{-30}$

For compactness, we refer to the CDF based on comparing the rMTF after the 4 Hz masker against the rMTF after the unmodulated masker; for example, rMTF (4 Hz) versus rMTF (0 Hz) as "4 vs 0," etc. Median values are based on all the data used to generate the CDFs and include all tested modulation frequencies. p-values in the rightmost column reflect Wilcoxon rank-sum tests on the distribution of response rate differences for all difference functions and all modulation frequencies. Significant results are shown in black and nonsignificant results are shown in gray.

Table 2. Results of analyses comparing the loci of maximal suppression for pairs of CDFs

CDF$_1$	CDF$_2$	COM$_1$ (Hz)	COM$_2$ (Hz)	p-value
4 vs 0	10 vs 0	15.96	15.51	0.4754
4 vs 0	32 vs 0	15.96	34.46	0.0037
4 vs 0	96 vs 0	15.96	72.27	<0.0001
10 vs 0	32 vs 0	15.51	34.46	0.0002
10 vs 0	96 vs 0	15.51	72.27	<0.0001
32 vs 0	96 vs 0	34.46	72.27	<0.0001
10 vs 4	32 vs 4	45.39	70.79	0.0097
10 vs 4	96 vs 4	45.39	119.32	0.0024
32 vs 4	96 vs 4	70.79	119.32	0.0008
32 vs 10	96 vs 10	86.03	128.73	0.0008

Conventions for referring to the CDFs are the same as in Table 2. Values for the COM of each CDF are middle columns. p-values in the rightmost column reflect results of a permutation test comparing actual disparities in the COMs with simulated disparities based on iterated random data mixtures (see Materials and Methods). Significant results are shown in black and nonsignificant results are shown in gray.

these analyses are collected in Table 2. With the exception of the comparisons of 4 Hz and 10 Hz against the unmodulated baseline (top row), CDFs with higher-modulation-frequency maskers were characterized by higher COMs, indicating that the locus of suppression also shifted to higher modulation frequencies. This was true of the comparisons between 96 Hz and 32 Hz maskers, demonstrating that, although the total magnitude of suppression did not differ (see above), the distribution of suppression across modulation frequency differed significantly.

Modulation-frequency-specific effects on firing rate were statistically robust

The results described above do not depend on our choice to normalize the data by the maximum value from either rMTF when computing the difference functions. We also analyzed computed difference functions by subtracting the rMTFs associated with different maskers and normalizing each point (i.e., modulation frequency) by the sum of the SEs (across trials; $n = 20$) for each rMTF. This procedure minimizes the contribution of recordings in which the trial-to-trial variability was high relative to the CDF structure. This procedure yielded essentially identical results when analyzed for total suppression, tuning, and the locus of tuning (data not shown).

We also verified that modulation-frequency suppression was a feature of the most robustly synchronized responses in our data sample. To do so, we divided the data used to generate the CDF into halves based on whether the VS averaged over the vsMTFs for each comparison fell above or below the population median and then generated two CDFs. In every case, the shapes of the CDFs indicated that the magnitude and tuning of suppression was greater for data associated with better synchronization (data not shown). We determined whether these differences were statistically significant by comparing the differences in the magnitude and tuning of suppression when segregating the difference functions by the degree of synchronization against CDFs comprised by random mixtures of difference functions. Magnitudes were significantly ($p < 0.0001$) greater for the better synchronized responses for the 10 Hz–4 Hz, 32 Hz–4 Hz, 32 Hz–10 Hz, and 96 Hz–4 Hz CDFs, with trends ($p < 0.025$) for the remainder. Tuning, quantified by the variance of the CDF, was significantly ($p < 0.0001$) greater for the 96 Hz–4 Hz and 96 Hz–32 Hz CDFs, with trends ($p < 0.05$) for the 32 Hz–4 Hz and 96 Hz–10 Hz CDFs. These results confirm that modulation-frequency-specific contextual modulation is more prominent among recording sites that better synchronize to SAM signals.

Given the length of the experiments, another important concern was the possibility that changes in firing rates reflect changes in responsiveness over time ("drift"), rather than genuine contextual modulation effected by the maskers. As in prior work (Malone et al., 2013), we sorted the 20 trials for each tested modulation frequency chronologically into 10 "early" and 10 "late" trials. Data obtained in the second half of the first run and the first half of the second run were obtained ~13 min more closely in time than data from the first half of the first run and the second half of the second run. Median values for the SIs between rMTFs obtained in different contexts (e.g., after the 4 Hz masker versus the 32 Hz masker) did not differ significantly when using the more proximal trials versus the more distal trials (0.83 vs 0.83; Wilcoxon rank-sum; $p > 0.45$), indicating that recording conditions were quite stable.

As a further check on the possibility that response drift could account for our results, we recomputed the CDFs after compensating for changes in the spontaneous rates, which should also be sensitive to changes in overall responsiveness over time. Specifically, we constructed difference functions based on the spontaneous rates measured after the end of the SAM signal on each trial and subtracted these from the rMTF-derived difference functions before averaging them to generate the CDFs. The drift-corrected CDFs were very similar to the actual CDFs and clearly indicated that response drift could not account for our results (data not shown). It should also be noted that, given the pseudorandom presentation of the SAM signals within each masker block (Fig. 1), it is exceedingly unlikely that response drift could ever produce the tuned adaptation effects we report here.

Masking effects were much greater within 500 ms of the masker offset but persisted throughout the duration of the SAM signals

We evaluated the time course of the contextual modulation induced by the maskers by computing two sets of CDFs based on either the first or second 500 ms of the responses to the probes. The results of this analysis are shown in Figure 6 (Fig. 6a–d corresponds to Fig. 4a–d; Fig. 6e–j corresponds to Fig. 5a–f). As is evident, over time, the CDFs appear to converge on the flat functions that one would expect if there were no effects of modulation
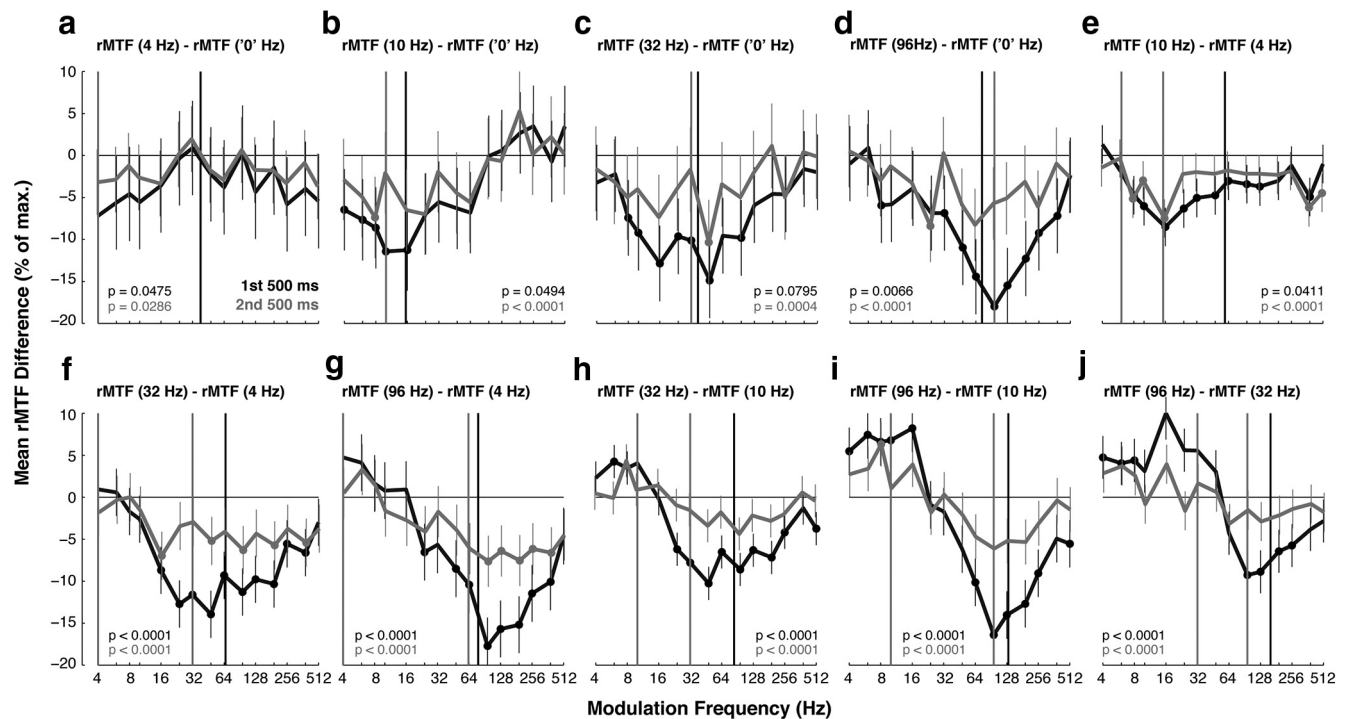
Figure 6. Figure conventions are similar to those used in Figures 4 and 5. The identities of the rMTFs and the order of subtraction used to produce the depicted CDFs are indicated by the titles appearing above panels *a–j*. The black curves represent CDFs calculated from the initial 500 ms of the SAM signal and the gray curves represent CDFs calculated from the terminal 500 ms of the SAM signal. The significance values in black indicate the results of a permutation test for differences in the variances of the black and gray curves (see Results). The significance values in gray indicate the results of a permutation test for persistent CDF tuning (see Results).

frequency context. To verify that the curves based on responses during the first (black curve) and second (gray curve) halves of the probes differed significantly, we performed permutation tests to demonstrate that the difference in the variances of the CDFs for the two response epochs were greater than could be explained by random mixtures of difference functions from both epochs. The significance values obtained from this analysis are indicated in black in Figure 6. Generally speaking, CDFs that showed the largest suppressive effects also exhibited significant reductions in CDF tuning over the course of 500 ms. We also tested whether the residual CDF tuning measured in the latter epoch was significant using the same methods applied to the data in Figures 4 and 5. Results of this analysis are indicated in gray. Contextual modulation effects in the first 500 ms after masker offset were strongest. However, most CDFs based on responses >500 ms from masker offset exhibit attenuated but similarly tuned contextual effects on cortical firing rates.

Modulation-frequency-specific adaptation effects on neural response synchronization were relatively modest
We applied the same analyses to the results for the vsMTFs that we previously described for the rMTFs. For vsMTFs, however, the units of the analysis are VS values bounded from 0 to 1 and no normalization was applied. Because VS measures the concentration of spike times within a particular modulation phase rather than rate, we refer to reductions in VS as "desynchronization" rather than "suppression." On a site by site basis, we found that the prevalence of significant contextual modulation of vsMTFs was lower than for rMTFs. Using a permutation test on the SIs, ~30% (449/1480) of the sites exhibited significant differences compared with 56% for the rMTFs. When we normalized the

vsMTFs by their sums before the analysis, 17% (250/1480) of the sites remained significant compared with 21% for the rMTFs.

Figure 7 shows the set of CDFs based on the vsMTFs organized similarly to Figure 6. We applied the same analyses of total suppression, tuning, and the locus of tuning we had used for the rMTF-derived CDFs to the vsMTF-derived CDFs. Presentation of the SAM maskers produced significant but modest desynchronization of the responses to the SAM signals. The median values for all points comprising the CDFs were -0.008, -0.006, -0.0001, and -0.0109 for the 4, 10, 32, and 96 Hz maskers, respectively, relative to the unmodulated maskers. These medians were significantly different from 0 for the 10 Hz ($p = 0.0041$) and 96 Hz ($p = 0.0003$) maskers, but not for the 4 Hz ($p = 0.0234$) or 32 Hz ($p = 0.2376$) maskers. However, there were no significant differences between the median values for any comparable CDF pairs ($n = 6$; $p > 0.09$ in all cases), indicating that higher-modulation-frequency maskers did not produce greater desynchronization despite producing greater firing rate suppression. The median values across all of the difference functions for modulated maskers at different modulation frequencies were similarly near 0 (10 Hz–4 Hz: 0.0002; 32 Hz–4 Hz: -0.0010; 96 Hz–4 Hz: -0.0027; 32 Hz–10 Hz: -0.0000; 96 Hz–10 Hz: -0.0029; 96 Hz–32 Hz: -0.0011). None of the medians differed significantly from 0 ($p > 0.04$ in all cases), nor were comparisons among pairs of medians significant ($p > 0.03$ in all cases; $p > 0.15$ in 10/11 cases).

As we had done for the rMTFs, we evaluated the tuning of contextually mediated desynchronization using Monte Carlo methods based on the variances of the CDFs. Significance values associated with this analysis are displayed in Figure 7. Relative to the CDFs for firing rate, significant deviations from 0 for partic-
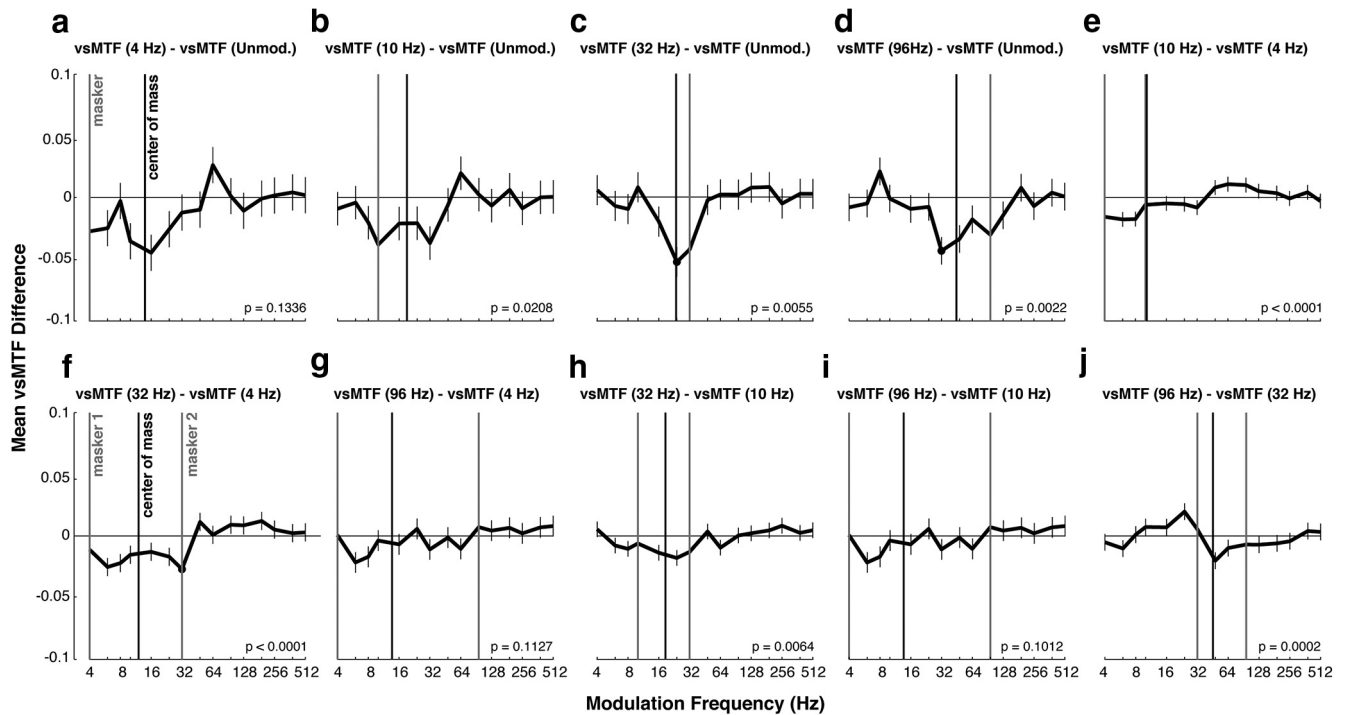
Figure 7. Figure conventions are similar to those used in Figure 4 except the ordinate represents VS rather than the proportional reduction in firing rate. The identities of the vsMTFs and the order of subtraction used to produce the depicted CDFs are indicated by the titles appearing above panels **a–j**. Significance values indicate the results of the permutation test of CDF tuning (see Materials and Methods).

ular modulation frequencies were comparatively rare and overall the curves exhibited relatively poorer tuning. Nevertheless, some CDFs indicated significant tuning, although some of the CDFs that exhibited the greatest rate suppression lacked significant desynchronization (e.g., 96 Hz–4 Hz and 96 Hz–10 Hz). Comparison of Figures 4 and 5 against Figure 7 suggests that the SAM maskers had greater effects on response rate than on response timing. We verified this by comparing the z-scores obtained when computing the permutation tests based on the CDF variances (see Materials and Methods). The results of this analysis are shown in Figure 8. For all CDFs, the z-scored variance was larger—sometimes by an order of magnitude—for firing rate relative to vector strength. Direct comparison of the z-scores indicated that the SAM maskers elicited greater response suppression than desynchronization ($p = 0.002$; Wilcoxon signed rank). Therefore, the prior presentation of the maskers tended to eliminate spikes such that synchrony with the modulation envelopes of the probes was relatively unperturbed.

Given the relative flatness of the vsMTF-derived CDFs, we expected that the orderly shifts in the locus of suppression that we reported for rMTF-derived CDFs would be less evident with respect to desynchronization. We performed a similar permutation test based on differences in the COMs for CDF pairs. Although the COM for the 96 Hz–0 Hz CDF was shifted to significantly ($p = 0.0074$) higher modulation frequencies compared with the 4 Hz–0 Hz CDF, none of the remaining comparisons involving the unmodulated baseline ($n = 5$) were significant ($p > 0.04$ in all cases). For CDFs based on two different modulated maskers, only the comparison between the 96 Hz–10 Hz and 32 Hz–10 Hz CDFs was significant ($p = 0.0027$; $p > 0.09$ for all other comparisons). Therefore, the evidence for modulation-frequency-specific desynchronization appears to be weaker than that for response suppression.
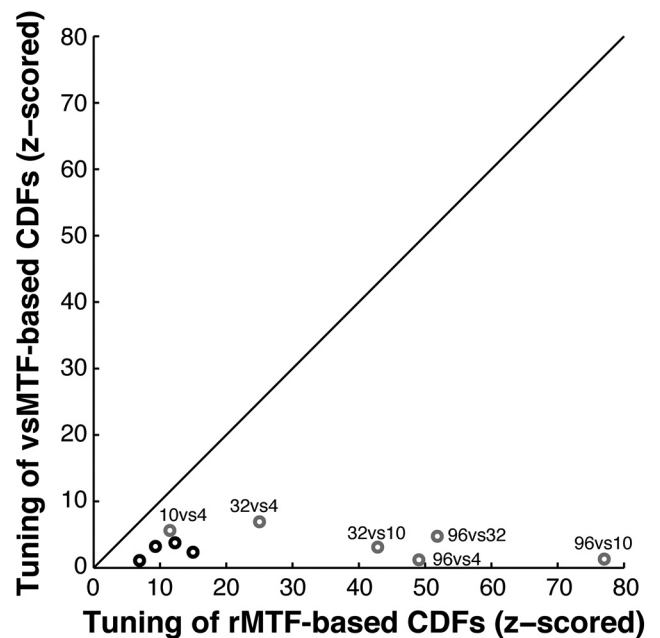

Figure 8. Scatterplot comparing the z-scores obtained when performing the permutation tests used to assign significance to the tuning of the CDFs based on rMTFs (abscissa) and vsMTFs (ordinate). Black circles indicate the z-scores for CDFs that include the unmodulated masker (Fig. 4). Gray circles indicate the z-scores for CDFs based on two different modulated maskers (Fig. 5). Identities for the latter are indicated adjacent to each circle.

Recent spiking history is a modest predictor of contextual effects for modulation frequency

If the presentation of the maskers produces the contextual modulation of the ensuing responses to the SAM signals, then the strength of a given site's response to the masker may predict the

magnitude of such effects. We evaluated this possibility at varying degrees of resolution.

First, we attempted to determine whether the genuine trend for greater suppression by modulated maskers was reflected in the distributions of firing rates elicited by the different maskers. Across all clusters, the median firing rates associated with the 0, 4, 10, 32, and 96 Hz maskers were 15.1, 29.9, 34.0, 28.2, and 29.6 Hz, respectively. The unmodulated ("0 Hz") maskers elicited significantly lower rates than the 4 modulated maskers (Wilcoxon rank-sum; $p < 10^{-9}$ in all cases). This comports with the fact that the 10, 32, and 96 Hz maskers elicited greater total suppression than the unmodulated maskers (see above), but does not explain why the 4 Hz masker did not. In a similar vein, the statistical trends suggesting higher firing rates for the 10 Hz masker relative to the 4 Hz ($p = 0.0220$) and 96 Hz ($p = 0.0190$) maskers cannot explain the fact that total suppression was consistently higher for the 32 and 96 Hz maskers relative to the 4 and 10 Hz maskers.

We also evaluated the ability of recent spiking history to predict contextual modulation for each modulation frequency in our data sample. Specifically, we investigated how effectively one could predict the (trial-averaged) difference in firing rates associated with a given modulation frequency (e.g., 24 Hz) by knowing the (trial-averaged) difference in firing rates elicited by the different maskers preceding it (e.g., 4 Hz vs 32 Hz). This correlation for all possible comparisons ($n = 22250$) was negative ($r = -0.28$; $p = 0$), as would be expected if higher firing rates during the masker correspond to lower firing rates during the subsequent probe stimulus. However, the predictions based only on masker-elicited firing rates accounted for a modest percentage of the total response variance. We found that this relationship was stronger when we limited the analysis to masker pairs that included the unmodulated masker ($r = -0.40$; $p < 10^{-142}$) and slightly weaker when both maskers were modulated ($r = -0.23$; $p < 10^{-214}$). It is possible that this difference reflects the greater disparity in firing rates across the modulated and unmodulated maskers described above.

The tuning of the contextual effects described above necessarily limits the explanatory power of recent spiking history. Because the same masker precedes each probe stimulus comprising the rMTF in a given context, differences in spiking history would predict a constant difference between two rMTFs obtained in different contexts, subject to variability for repeated trials. We would expect a correlation driving solely by spiking history to be maximal when the suppressive effects of the more effective masker are largest. We computed the average firing rate associated with each masker (across all presentations for a given rMTF) and then compared the difference in firing rates elicited by two different maskers against the single point of maximal suppression for each rMTF pair. In this case, the correlation strengthened only modestly ($r = -0.42$; $p < 10^{-65}$).

We repeated these analyses for pairs of vsMTFs obtained with different maskers. Because the maskers tended to be more effective in suppressing firing rates than in desynchronizing the spiking patterns, we expected the correlations to be weaker. In fact, there was a weak but significant positive correlation between the difference in firing rates associated with the makers and the VS measured during the SAM signals ($r = 0.05$; $p < 10^{-13}$); that is, a stronger response to the masker slightly improved synchronization to the subsequent SAM signal, suggesting that the suppression of the response was more likely to eliminate spikes occurring at modulation phases distinct from the mean phase.

Discussion

Our results clarify multiple aspects of cortical modulation masking. First, contextual rate suppression exhibited broad but significant tuning. Suppressive effects were concentrated at or near the masker modulation frequency. Second, the response rates elicited by the maskers were only modestly predictive of the degree of observed suppression. Third, the adaptation effects were largely confined to changes in firing rate—the synchronization of the remaining spikes with the SAM envelopes was essentially unaffected. This disjunction between rate and temporal AM encoding has implications for how cortical signals inform behavioral performance. Fourth, the magnitude of the suppression decreased as the time from the masker offset increased. Nevertheless, it remained significant 500–1000 ms after masker offset. Finally, we found that masker stimuli at the higher tested modulation rates (32 and 96 Hz vs 4 and 10 Hz) produced greater suppression overall.

Our finding that cortical neurons exhibit tuned modulation masking is consistent with psychoacoustical modulation-masking results. Tuning was broad in accordance with reported Q values (the ratio of the center frequency and bandwidth) of modeled modulation filters (Ewert and Dau, 2000; Ewert et al., 2002; Sek and Moore, 2002, 2003; Wojtczak and Viemeister, 2005; 0.35–2). However, psychophysical tMTFs are based on elevated thresholds for modulation detection, whereas our physiological rMTFs are based on 100% modulated stimuli. The perceived modulation depth of suprathreshold SAM stimuli can also be contextually reduced in forward-masking paradigms, although signals with depths at or near 100% were not affected appreciably (Wojtczak and Viemeister, 2003). Therefore, we must be cautious when comparing physiological evidence of rate suppression with psychoacoustic evidence of threshold changes for modulation detection. However, prior presentation of SAM maskers could reduce the detection rate of fully modulated (100%) SAM probes that are otherwise reliably detected to values near threshold (~75% correct; Wojtczak and Viemeister, 2005), suggesting a correspondence between physiological and psychophysical modulation masking. Modulation-frequency-specific rate suppression is broadly consistent with the operation of a modulation filterbank modeled on the band-pass MTFs described in the IC (Langner and Schreiner, 1988) and explicitly incorporated into computational models of modulation detection and masking (Dau et al., 1997; Jepsen et al., 2008).

Our results suggest that modulation masking may be instantiated in the forebrain. Bartlett and Wang (2005) presented a range of different modulation frequencies as forward-masking stimuli for a single (best) modulation frequency probe. They observed long-lasting suppression and facilitation in individual cortical neurons, with suppression dominant for maskers within an octave of the probe modulation frequency. Their results suggest that modulation context effects are idiosyncratic but exhibit commonalities that would yield something like the effects we report here when averaged across neurons. In contrast, Wojtczak et al. (2011) found that ~60% of IC neurons of awake rabbits did not exhibit rate modulation masking, suggesting that robust modulation masking emerges above the level of the IC. Modulation masking may emerge later in the auditory pathway than spectral masking simply because rate tuning for modulation frequency emerges later in the auditory pathway than rate tuning for spectral frequency (Joris and Yin, 1992; Malone and Semple, 2001; Joris et al., 2004; Nelson et al., 2009).

Rate tuning for a given stimulus feature could suffice for context-sensitive adaptation if the underlying biophysical mechanisms require only synaptic and/or spiking activity to be engaged. The distinction between synaptic and spiking activity is essential because it specifies whether contextual adaptation is caused by a neuron's synaptic input or spiking output. A central neuron's own spiking history reflects but cannot fully capture the synaptic and spiking histories of the many neurons comprising its signal pathway for a given stimulus. Auditory nerve fibers reduce their responses to probe tones in direct proportion to the strength of their response to the forward masker (Harris and Dallos, 1979; Relkin and Turner, 1988). In contrast, forward suppression in the IC (Malone et al., 2001; Nelson et al., 2009), medial geniculate body (Schreiner, 1981), and cortex (Calford and Semple, 1995; Brosch and Schreiner, 1997) is often not effectively explained by spiking history. The decoupling of spiking history and adaptation effects—including the weak correlations we found between responses to masker and probe SAM—may occur because the recorded neuron's spiking activity represents only a fraction of the relevant response history of the entire signal pathway.

When all probe stimuli share a common masker, models based solely on the spiking history must predict uniform suppression. Instead, we found that suppression was proportional to the similarity between the modulation frequencies of the masker and probe, suggesting that the adaptation of the probe response reflects the prior engagement of a subset of the neuron's inputs by the masker. When the masker and probe SAM differ sufficiently, however, inputs tuned to the probe are not adapted by the masker and respond more robustly to the probe. Interestingly, Scholes et al. (2011) observed that the predictive validity of spiking history increased with increasing probe level, suggesting that the recorded neuron's spiking history is more constraining when subsequently tasked with firing at higher rates. With respect to modulation filterbanks, the tuned response suppression we observe is not necessarily a reflection of the existence of labeled-line AM channels, but rather the biased convergence of partially overlapping AM-tuned inputs at multiple stages of the auditory pathway up to and including the cortex. This arrangement also suggests a relative enhancement of the cortical firing rate representation of changes in modulation statistics due to the recruitment of "fresh" afferent activity by such changes (May and Tiitinen, 2010).

Modulation-frequency-specific forward suppression of average firing rate cooccurred with modest effects on cortical synchronization (Figs. 7,8). The functional implications of this disjunction depend on how the responses of neurons in the auditory core are decoded by downstream neural populations governing behavior. If physiological rate suppression underlies the psychoacoustic phenomenon of modulation masking, then the cortical decoding strategy used by human listeners must reflect the average rate rather than the cycle-by-cycle rate fluctuations. Cortical firing rates increase monotonically with increasing modulation depth (Malone et al., 2010), suggesting that modulation is reliably associated with higher response rates. Nevertheless, spike timing information is demonstrably better at decoding AM and FM frequency from cortical spike trains out to 100 Hz, whereas average firing rate information is consistently poor (Malone et al., 2007, 2014). Therefore, it is puzzling that such information would go unused (Lemus et al., 2009; Dong et al., 2011), but consistent with the finding that firing rate is more predictive of psychophysically reported AM detection than is phase locking (Niwa et al., 2012). The mnemonic demands of the particular psychophysical paradigms (e.g., three-interval forced-choice;

Wojtczak and Viemeister, 2005) may be relevant here given the limited auditory working memory of humans (Bigelow and Poremba, 2014) and nonhuman primates (Scott et al., 2012).

It is also likely that the perceptual strategies for detecting modulation vary with modulation frequency (Gutschalk et al., 2008; Edwards and Chang, 2013). Maskers within the "fluctuation range" critical for speech (4 and 10 Hz) produced significantly less rate suppression than those in the "roughness" range (32 and 96 Hz). Reduced rate suppression for very low modulation frequencies could reflect partial recovery from adaptation within individual modulation cycles. Macaque monkeys trained to detect modulated tones amid a background of modulated noise enjoyed an advantage based on a relative modulation phase shift (180°) at 10 Hz, but not at or >20 Hz (Bohlen et al., 2014). Even modeled modulation filters tuned <10 Hz explicitly retain information about modulation phase (Jepsen et al., 2008). Unfortunately, the most methodologically similar psychoacoustic demonstration of modulation masking (Wojtczak and Viemeister, 2005) used modulation rates (20, 40, and 80 Hz) beyond the synchronization limits of most cortical neurons (Liang et al., 2002; Malone et al., 2007), whereas the use of very low (e.g., 4 Hz) modulation frequencies remains confined to simultaneous masking paradigms (Bacon and Grantham, 1989; Ewert and Dau, 2004). Further work will be necessary to characterize how the perceptual decoding strategies used by different listeners (Niwa et al., 2012, 2013, 2015) and the time constants of the biophysical mechanisms responsible for adaptation interact with modulation frequency and shape AM perception in general (Riecke et al., 2014).

The modulation filterbank hypothesis has been generally successful in accounting for the psychophysical modulation detection and masking results it was created to explain (Dau et al., 1997). Xiang et al. (2013) observed nonlinear interactions among putative modulation filters (i.e., sum and difference frequencies) using MEG in the context of simultaneous masking. Our results are best understood as explaining why modulation filterbank models succeed based on modulation-frequency-specific rate suppression in cortical neurons. Nevertheless, physiological MTFs exhibit important differences from idealized modulation filterbanks (Kay, 1982), including a lack of invariance for stimulus parameters that do not directly affect stimulus periodicity (IC: Krebs et al., 2008; Zheng and Escabi, 2008; Cortex: Malone et al., 2007, 2010, 2013, 2014). Despite the fact that central auditory neurons fail to correspond to idealized modulation filters in important ways, our findings suggest that averaging cortical responses over a suitably large population produces physiological modulation-masking profiles that effectively capture essential features of psychoacoustic-masking results such as tuning, specificity, and persistence. Cortical evidence of rate suppression in the absence of similar evidence in the midbrain (Wojtczak et al., 2011) is consistent with the emergence of modulation-masking phenomena in the auditory forebrain.

References

Bacon SP, Grantham DW (1989) Modulation masking: effects of modulation frequency, depth, and phase. J Acoust Soc Am 85:2575–2580. CrossRef Medline

Bartlett EL, Wang X (2005) Long-lasting modulation by stimulus context in primate auditory cortex. J Neurophysiol 94:83–104. CrossRef Medline

Bieser A, Müller-Preuss P (1996) Auditory responsive cortex in the squirrel monkey: neural responses to amplitude-modulated sounds. Exp Brain Res 108:273–284. Medline

Bigelow J, Poremba A (2014) Achilles' ear? Inferior human short-term and

recognition memory in the auditory modality. PLoS One 9:e89914. CrossRef Medline

Bohlen P, Dylla M, Timms C, Ramachandran R (2014) Detection of modulated tones in modulated noise by non-human primates. J Assoc Res Otolaryngol 15:801–821. CrossRef Medline

Brosch M, Schreiner CE (1997) Time course of forward masking tuning curves in cat primary auditory cortex. J Neurophysiol 77:923–943. Medline

Calford MB, Semple MN (1995) Monaural inhibition in cat auditory cortex. J Neurophysiol 73:1876–1891. Medline

Cazals Y, Pelizzone M, Saudan O, Boex C (1994) Low-pass filtering in amplitude modulation detection associated with vowel and consonant identification in subjects with cochlear implants. J Acoust Soc Am 96:2048–2054. CrossRef Medline

Cheung SW, Bedenbaugh PH, Nagarajan SS, Schreiner CE (2001) Functional organization of squirrel monkey primary auditory cortex: responses to pure tones. J Neurophysiol 85:1732–1749. Medline

Creutzfeldt O, Hellweg FC, Schreiner C (1980) Thalamocortical transformation of responses to complex auditory stimuli. Exp Brain Res 39:87–104. Medline

Dau T, Kollmeier B, Kohlrausch A (1997) Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers. J Acoust Soc Am 102:2892–2905.

Dolležal LV, Beutelmann R, Klump GM (2012) Stream segregation in the perception of sinusoidally amplitude-modulated tones. PLoS One 7:e43615. doi: 10.1371/journal.pone.0043615. CrossRef Medline

Dong C, Qin L, Liu Y, Zhang X, Sato Y (2011) Neural responses in the primary auditory cortex of freely behaving cats while discriminating fast and slow click-trains. PLoS One 6:e25895. CrossRef Medline

Edwards E, Chang EF (2013) Syllabic (∼2–5 Hz) and fluctuation (∼1–10 Hz) ranges in speech and auditory processing. Hear Res 305:113–134. CrossRef Medline

Ewert SD, Dau T (2000) Characterizing frequency selectivity for envelope fluctuations. J Acoust Soc Am 108:1181–1196. Medline

Ewert SD, Dau T (2004) External and internal limitations in amplitude-modulation processing. J Acoust Soc Am 116:478–490. CrossRef Medline

Ewert SD, Verhey JL, Dau T (2002) Spectro-temporal processing in the envelope-frequency domain. J Acoust Soc Am 112:2921–2931. CrossRef Medline

Fu QJ (2002) Temporal processing and speech recognition in cochlear implant users. Neuroreport 13:1635–1639. CrossRef Medline

Goldberg JM, Brown PB (1969) Response of binaural neurons of dog superior olivary complex to dichotic total stimuli: some physiological mechanisms of sound localization. J Neurophysiol 32:613–636. Medline

Grimault N, Bacon SP, Micheyl C (2002) Auditory stream segregation on the basis of amplitude-modulation rate. J Acoust Soc Am 111:1340–1348. CrossRef Medline

Gutschalk A, Micheyl C, Oxenham AJ (2008) Neural correlates of auditory perceptual awareness under informational masking. PLoS Biol 6:e138. CrossRef Medline

Harris DM, Dallos P (1979) Forward masking of auditory nerve fiber responses. J Neurophysiol 42:1083–1107. Medline

Houtgast T (1989) Frequency selectivity in amplitude-modulation detection. J Acoust Soc Am 85:1676–1680. CrossRef Medline

Jepsen ML, Ewert SD, Dau T (2008) A computational model of human auditory signal processing and perception. J Acoust Soc Am 124:422–438. CrossRef Medline

Johnson JS, Yin P, O'Connor KN, Sutter ML (2012) Ability of primary auditory cortical neurons to detect amplitude modulation with rate and temporal codes: neurometric analysis. J Neurophysiol 107:3325–3341. CrossRef Medline

Joris PX, Yin TC (1992) Responses to amplitude-modulated tones in the auditory nerve of the cat. J Acoust Soc Am 91:215–232. CrossRef Medline

Joris PX, Schreiner CE, Rees A (2004) Neural processing of amplitude modulated sounds. Physiol Rev 84:541–577. CrossRef Medline

Kay RH (1982) Hearing of modulation in sounds. Physiol Rev 62:894–975. Medline

Kay RH, Matthews DR (1971) Temporal specificity in human auditory conditioning by frequency-modulated tones. J Physiol 218:104P–106P. Medline

Kay RH, Matthews DR (1972) On the existence in human auditory pathways of channels selectively tuned to the modulation present in frequency-modulated tones. J Physiol 225:657–677. CrossRef Medline

Kohlrausch A, Fassel R, Dau T (2000) The influence of carrier level and frequency on modulation and beat-detection thresholds for sinusoidal carriers. J Acoust Soc Am 108:723–734. CrossRef Medline

Krebs B, Lesica NA, Grothe B (2008) The representation of amplitude modulations in the mammalian auditory midbrain. J Neurophysiol 100:1602–1609. CrossRef Medline

Krishna BS, Semple MN (2000) Auditory temporal processing: responses to sinusoidally amplitude-modulated tones in the inferior colliculus. J Neurophysiol 84:255–273. Medline

Langner G (1992) Periodicity coding in the auditory system. Hear Res 60:115–142. CrossRef Medline

Langner G, Schreiner CE (1988) Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. J Neurophysiol 60:1799–1822. Medline

Lemus L, Hernández A, Romo R (2009) Neural codes for perceptual discrimination of acoustic flutter in the primate auditory cortex. Proc Natl Acad Sci U S A 106:9471–9476. CrossRef Medline

Liang L, Lu T, Wang X (2002) Neural representations of sinusoidal amplitude and frequency modulations in the primary auditory cortex of awake primates. J Neurophysiol 87:2237–2261. Medline

Lu T, Liang L, Wang X (2001) Neural representations of temporally asymmetric stimuli in the auditory cortex of awake primates. J Neurophysiol 85:2364–2380. Medline

Malone BJ, Schreiner CE (2010) Coding of time varying sounds: envelope modulations. In: The Oxford handbook of auditory science: the auditory brain, Vol 2 (Rees A, Palmer A, eds), pp 123–128. Oxford: OUP.

Malone BJ, Semple MN (2001) Effects of auditory stimulus context on the representation of frequency in the gerbil inferior colliculus. J Neurophysiol 86:1113–1130. Medline

Malone BJ, Scott BH, Semple MN (2007) Dynamic amplitude coding in the auditory cortex of awake rhesus macaques. J Neurophysiol 98:1451–1474. CrossRef Medline

Malone BJ, Scott BH, Semple MN (2010) Temporal codes for amplitude contrast in auditory cortex. J Neurosci 30:767–784. CrossRef Medline

Malone BJ, Beitel RE, Vollmer M, Heiser MA, Schreiner CE (2013) Spectral context affects temporal processing in awake auditory cortex. J Neurosci 33:9431–9450. CrossRef Medline

Malone BJ, Scott BH, Semple MN (2014) Encoding frequency contrast in primate auditory cortex. J Neurophysiol 111:2244–2263. CrossRef Medline

Mardia K, Jupp P (2000) Directional statistics. New York: Wiley.

May PJ, Tiitinen H (2010) Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. Psychophysiology 47:66–122. CrossRef Medline

McDermott JH, Simoncelli EP (2011) Sound texture perception via statistics of the auditory periphery: evidence from sound synthesis. Neuron 71:926–940. CrossRef Medline

Nelson PC, Smith ZM, Young ED (2009) Wide-dynamic-range forward suppression in marmoset inferior colliculus neurons is generated centrally and accounts for perceptual masking. J Neurosci 29:2553–2562. CrossRef Medline

Niwa M, Johnson JS, O'Connor KN, Sutter ML (2012) Activity related to perceptual judgment and action in primary auditory cortex. J Neurosci 32:3193–3210. CrossRef Medline

Niwa M, Johnson JS, O'Connor KN, Sutter ML (2013) Differences between primary auditory cortex and auditory belt related to encoding and choice for AM sounds. J Neurosci 33:8378–8395. CrossRef Medline

Niwa M, O'Connor KN, Engall E, Johnson JS, Sutter ML (2015) Hierarchical effects of task-engagement on amplitude modulation encoding in auditory cortex. J Neurophysiol 113:307–327. Medline

Relkin EM, Turner CW (1988) A reexamination of forward masking in the auditory nerve. J Acoust Soc Am 84:584–591. CrossRef Medline

Riecke L, Scharke W, Valente G, Gutschalk A (2014) Sustained selective attention to competing amplitude-modulations in human auditory cortex. PLoS One 9:e108045. CrossRef Medline

Scholes C, Palmer AR, Sumner CJ (2011) Forward suppression in the auditory cortex is frequency-specific. Eur J Neurosci 33:1240–1251. CrossRef Medline

Schreiner CE (1981) Post-stimulatory effects in the medial geniculate body

of guinea pigs. In: Neuronal mechanisms of hearing (Syka J, Aitkin L, eds), pp 191–196. New York: Plenum.

Schreiner CE, Urbas JV (1986) Representation of amplitude modulation in the auditory cortex of the cat. I. The anterior auditory field (AAF). Hear Res 21:227–241. CrossRef Medline

Schreiner CE, Urbas JV (1988) Representation of amplitude modulation in the auditory cortex of the cat. II. Comparison between cortical fields. Hear Res 32:49–63. CrossRef Medline

Scott BH, Malone BJ, Semple MN (2011) Transformation of temporal processing across auditory cortex of awake macaques. J Neurophysiol 105: 712–730. CrossRef Medline

Scott BH, Mishkin M, Yin P (2012) Monkeys have a limited form of short-term memory in audition. Proc Natl Acad Sci U S A 109:12237–12241. CrossRef Medline

Sek A, Moore BC (2002) Mechanisms of modulation gap detection. J Acoust Soc Am 111:2783–2792. CrossRef Medline

Sek A, Moore BC (2003) Testing the concept of a modulation filter bank: the audibility of component modulation and detection of phase change in three-component modulators. J Acoust Soc Am 113:2801–2811. CrossRef Medline

Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M (1995) Speech recognition with primary temporal cues. Science 270:303–304. CrossRef Medline

Viemeister NF (1979) Temporal modulation transfer functions based upon modulation thresholds. J Acoust Soc Am 66:1364–1380. CrossRef Medline

Wojtczak M, Viemeister NF (2003) Suprathreshold effects of adaptation produced by amplitude modulation. J Acoust Soc Am 114:991–997. CrossRef Medline

Wojtczak M, Viemeister NF (2005) Forward masking of amplitude modulation: basic characteristics. J Acoust Soc Am 118:3198–3210. CrossRef Medline

Wojtczak M, Nelson PC, Viemeister NF, Carney LH (2011) Forward masking in the amplitude-modulation domain for tone carriers: psychophysical results and physiological correlates. J Assoc Res Otolaryngol 12:361–373. CrossRef Medline

Won JH, Drennan WR, Nie K, Jameyson EM, Rubinstein JT (2011) Acoustic temporal modulation detection and speech perception in cochlear implant listeners. J Acoust Soc Am 130:376–388. CrossRef Medline

Xiang J, Poeppel D, Simon JZ (2013) Physiological evidence for auditory modulation filterbanks: cortical responses to concurrent modulations. J Acoust Soc Am 133:EL7–EL12. CrossRef Medline

Zheng Y, Escabí MA (2008) Distinct roles for onset and sustained activity in the neuronal code for temporal periodicity and acoustic envelope shape. J Neurosci 28:14230–14244. CrossRef Medline

Zhou Y, Wang X (2014) Spatially extended forward suppression in primate auditory cortex. Eur J Neurosci 39:919–933. CrossRef Medline