UC Davis
UC Davis Previously Published Works

Title

Active Engagement Improves Primary Auditory Cortical Neurons' Ability to Discriminate Temporal Modulation

Permalink

https://escholarship.org/uc/item/2b44r4fp

Journal

Journal of Neuroscience, 32(27)

ISSN

0270-6474

Authors

Niwa, Mamiko
Johnson, Jeffrey S
O'Connor, Kevin N
et al.

Publication Date

2012-07-04

DOI

10.1523/jneurosci.5832-11.2012

Peer reviewed

Behavioral/Systems/Cognitive

Active Engagement Improves Primary Auditory Cortical Neurons' Ability to Discriminate Temporal Modulation

Mamiko Niwa, Jeffrey S. Johnson, Kevin N. O'Connor, and Mitchell L. Sutter

Center for Neuroscience and Department of Neurobiology, Physiology, and Behavior, University of California, Davis, Davis, California 95618

The effect of attention on single neuron responses in the auditory system is unresolved. We found that when monkeys discriminated temporally amplitude modulated (AM) from unmodulated sounds, primary auditory cortical (A1) neurons better discriminated those sounds than when the monkeys were not discriminating them. This was observed for both average firing rate and vector strength (VS), a measure of how well neurons temporally follow the stimulus' temporal modulation. When data were separated by nonsynchronized and synchronized responses, the firing rate of nonsynchronized responses best distinguished AM-noise from unmodulated noise, followed by VS for synchronized responses, with firing rate for synchronized neurons providing the poorest AM discrimination. Firing rate-based AM discrimination for synchronized neurons, however, improved most with task engagement, showing that the least sensitive code in the passive condition improves the most with task engagement. Rate coding improved due to larger increases in absolute firing rate at higher modulation depths than for lower depths and unmodulated sounds. Relative to spontaneous activity (which increased with engagement), the response to unmodulated sounds decreased substantially. The temporal coding improvement—responses more precisely temporally following a stimulus when animals were required to attend to it— expands the framework of possible mechanisms of attention to include increasing temporal precision of stimulus following. These findings provide a crucial step to understanding the coding of temporal modulation and support a model in which rate and temporal coding work in parallel, permitting a multiplexed code for temporal modulation, and for a complementary representation of rate and temporal coding.

Introduction

Processing temporal modulation is vital to interpreting sounds. Speech, for example, is laden with meaningful temporal cues, such as amplitude modulations (AMs), which serve as information-bearing parameters in speech recognition (Van Tasell et al., 1987; Shannon et al., 1995). AM is also a powerful cue for segregating sound sources in complex listening environments (Bregman et al., 1990; Grimault et al., 2002). Auditory cortical lesions, including those restricted to the primary auditory cortex (A1), have demonstrated that auditory cortex is necessary for temporal and language/vocal communication processing (Heffner and Heffner, 1986; Heffner and Heffner, 1989; Fitch et al., 1994; Griffiths et al., 1997). However, demonstrations of close links between the perception of temporal sound features and auditory cortical activity have proven elusive, although recently activity associated with perceptual choices in A1 has been reported using choice probability analysis (Niwa et al., 2012). In

anesthetized and awake nonbehaving animals, auditory cortical neurons may represent AM-related sound parameters by changes in average firing rate (rate code) and/or in temporal firing patterns (temporal code) (Lu et al., 2001; Liang et al., 2002; Schnupp et al., 2006; Malone et al., 2007; Kajikawa et al., 2008; Walker et al., 2008; Bizley et al., 2010; Imaizumi et al., 2010; Malone et al., 2010; Rosen et al., 2010). AM can be coded by the temporal pattern of activity directly mimicking the temporal pattern of the stimulus (phase-locking). It is known that neurons throughout the auditory system phase lock well (Review in Joris et al., 2004) and that auditory cortex extracts temporal speech features by tracking the temporal envelope (Steinschneider et al., 1980; Steinschneider et al., 1994; Eggermont, 1995; Schreiner, 1998; Steinschneider et al., 2005; Engineer et al., 2008). Because phase-locked coding is so fundamental to the auditory system, this system is ideal to investigate how behavioral state modulates temporal patterns of activity.

The activity of auditory cortical neurons in behaving animals depends on an animal's behavioral state and "attentiveness" (Hubel et al., 1959; Grady et al., 1997; Otazu et al., 2009). In addition, auditory task engagement has been shown to change A1 neuron response properties compared to when animals are passively presented with the same sounds. These changes—e.g., facilitative frequency tuning changes in tone detection (Miller et al., 1972; Fritz et al., 2003) and increased sharpness of spatial tuning in sound localization (Lee and Middlebrooks, 2011)— could be used to improve behavioral performance. In this paper we measured the ability of neurons to distinguish an AM sound

from its unmodulated noise carrier—the same discrimination the animals performed—both when the animals were performing the discrimination and when they were sitting passively. Each neuron's ability was assessed with both rate and phase-locked codes. We found that during behavioral discrimination, neural discriminability of AM improved relative to passive listening based on both rate and phase-locked codes, suggesting that the accuracy of representation in both codes can be modulated by behavioral state.

Materials and Methods

Subjects
Data were from the right hemispheres of two female (monkeys V and W) and one male (monkey X) adult rhesus monkeys (*Macaca mulatta*), each weighing 6–11 kg. Monkey initials are consistent across all publications from this laboratory. All procedures conformed to the United States Public Health Service policy on experimental animal care and were approved by the University of California, Davis animal care and use committee.

Acoustic stimuli
Stimuli were 800 ms sinusoidally AM broadband noise bursts (modulation frequencies: 2.5, 5, 10, 15, 20, 30, 60, 120, 250, 500, and 1000 Hz; modulation depths: 6, 16, 28, 40, 60, 80, and 100%) and unmodulated (0% modulation) broadband noise. The broadband noise carrier was "frozen"; the same random number sequence was used as a noise carrier sample for all stimuli. Sound generation has been described previously (O'Connor et al., 2011). Briefly, the sound signals were created in MatLab (The MathWorks) and generated using a D/A converter (model 1401; Cambridge Electronic Design). They were then passed through a programmable (PA5; TDT Systems) and a passive (LAT-45; Leader) attenuator, amplified (MPA-200; Radio Shack), and delivered to a speaker. Two different recording set-ups were used. One had a PA-110 (Radio Shack) speaker 1.5 m from the subject at its ear level. The other had an Optimus Pro-7AV (Radio Shack) positioned 0.8 m in front of the subject at its ear level. The sound was generated at a sampling rate of 100 kHz and had 5 ms cosine-ramped onsets and offsets. Intensity was calibrated with a sound-level meter (model 2231; Bruel and Kjaer) to 63 dB sound pressure level for all sounds at the outer ears.

Behavioral task
The behaving condition was discriminating AM noise from unmodulated noise (i.e., a modulation detection task, detecting whether the stimulus was amplitude modulated). The monkeys were trained to initiate a trial by pressing and holding down a lever. A trial consisted of two 800 ms sounds separated by a 400 ms interstimulus interval. The first (standard) sound was an unmodulated noise, and the second (test) sound was either unmodulated (nontarget) or an AM noise (target). Target stimuli had a fixed modulation frequency [at the multiple unit's best modulation frequency, tested from 2.5 to 1000 Hz; see Physiological recording for details of best modulation frequency (BMF) determination] during a recording session and modulation depths of 6, 16, 28, 40, 60, 80, and 100%. Subjects were trained to respond to AM targets by releasing the lever during an 800 ms response window following the offset of the second sound. When the second sound was unmodulated (0% depth), the subjects were required to hold down the lever for the entire response window. The macaques were rewarded with juice or water for both hits (a lever release for target trials) and correct rejections (holding down the lever for nontarget trials). Animals were notified on incorrect responses (misses and false alarms; not releasing the lever on target trials and releasing the lever on nontarget trials, respectively) by the offset of an incandescent light placed in front of them. False alarms were also followed by a time-out period of 15–60 s.

Behavioral training
Monkeys W and V went through the training described below. Monkey X underwent similar training but was previously used in an auditory induction experiment (Petkov et al., 2003).

After training to sit quietly in an "acoustically transparent" primate chair, the animals were taught to depress and release a lever for liquid reward using standard operant shaping techniques. They then were trained to depress the lever to initiate presentation of a brief (100 ms) AM noise (target) and were rewarded for releasing the lever after AM offset. The response limit was initially 10 s after offset and was decreased to 800 ms over the course of several sessions. Concurrently, the delay between initial lever depression and sound presentation was increased to 1 s. Next, two low-intensity 100 ms unmodulated noise bursts were introduced in the 1 s (previously silent) period between lever depression and AM presentation, with 100 ms silent intersound intervals (ISIs) among the three sounds. The intensity of the unmodulated noise was gradually increased until it was equal to that of the AM. Subsequently, ISIs of 200, 400, and 800 ms replaced the fixed 100 ms ISI, the duration of the stimulus was extended, and the number of possible pre-AM noise bursts was changed from two bursts to two, three, or four bursts. At this point, standard only (nontarget) trials were also introduced. This transition took several sessions. The time needed for asymptotic performance varied across subjects from 9 months to 1.5 years; therefore, these animals were highly trained by the time recording experiments began. During physiological recording sessions, all trials comprised one standard (unmodulated 800 ms noise), followed by a 400 ms ISI, and then a test stimulus (AM or unmodulated 800 ms noise), so that more data could be collected for AM relative to unmodulated noise. After this training, the animals were informed by cue light when blocks of behaving or passive stimulation conditions were to occur.

Physiological recording
The physiological recording procedures were similar to those previously described (O'Connor et al., 2010). Briefly, each monkey was chronically implanted with a titanium head holder and a CILUX recording chamber (Crist Instrument) placed over the parietal cortex to allow for the near vertical access to A1. A plastic grid with 27-gauge holes was placed on the recording chamber. The grid held a stainless steel, transdural guide tube that could be inserted throughout a 15 mm × 15 mm area of the brain at 1 mm intervals. A high-impedance tungsten microelectrode was inserted through the guide tube and lowered into A1 by a hydraulic microdrive. All recordings were made while the monkey sat, head restrained, in an "acoustically transparent" primate chair in a double-walled, sound-attenuated, foam-lined booth (IAC: 9.5′ × 10.5′ × 6.5′ or 4′ × 3′ × 6.5′).

Electrophysiological signals were amplified and filtered (0.3–10 kHz; AM Systems model 1800 and Krohn-Hite model 3382), passed to a computer with an A/D converter (sampling rate, 50 kHz; CED model 1401), and saved to hard disk along with the time stamps of all other relevant events for later analysis. Action potentials were sorted and assigned to individual neurons [single units (SUs)] online, and then refined offline using the waveform-matching algorithm by SPIKE2 (CED). Multiple units (MUs) were clear spiking activity collected above the background level.

The BMF was determined at each recording site's MU by presenting AM noise (modulation frequency: 2.5, 5, 10, 15, 20, 30, 60, 120, 250, 500, and 1000 Hz; modulation depth, 100%) and unmodulated noise. Then receiver operating characteristic (ROC) areas for AM at each modulation frequency were calculated based on firing rate and vector strength (VS) by comparing responses to AM noise with responses to unmodulated noise (see ROC analysis). BMF_{VS} was the modulation frequency with greatest VS-based ROC area, and BMF_{SC} was the frequency with rate-based ROC area most deviant from 0.5. When BMF_{VS} and BMF_{SC} were different, the selection of BMF_{VS} or BMF_{SC} was alternated from day to day to test equivalently sized samples at both modulation frequencies. After the BMF was selected, AM sensitivity was determined (1) while animals performed the AM detection task (behaving condition, target modulation frequency at the BMF, depths from 6 to 100%) and (2) while they were presented with the same stimuli and received randomly timed liquid rewards for sitting quietly (passive condition). For monkeys V and W, the order of behaving and passive experiments was alternated from recording day to day. For monkey X, behaving condition was always followed by passive condition.

Data analysis

Neural ROC analysis

Signal detection theory-based ROC analysis (Green and Swets, 1974; Relkin and Pelli, 1987; Britten et al., 1992; Spitzer et al., 2004; Scott et al., 2007) was used to quantify how well a neuron discriminates AM from its unmodulated carrier. First, a measure (e.g., firing rate) of the unit's responses was calculated on each trial. This was done for both AM sound (signal) and unmodulated noise (noise). We only analyzed unmodulated sounds that were in the second sound position (test sound) of the sequence. This is because the response to an identical sound in a two-sound sequence may differ, depending on whether it was first or second (Brosch and Schreiner, 1997; Werner-Reiss et al., 2006). Therefore, to allow fair comparisons to the targets that also were in the second position, we only analyzed unmodulated sounds presented in the same temporal position. Then the neural measure obtained in repeated trials was plotted into probability distributions for the AM at a given depth and the unmodulated noise. From these two probability distributions, we determined the proportion of trials in which neural response to AM exceeded a given criterion level. The proportion of trials in which neural response to the AM exceeded the criterion level is directly comparable to hit rate. The proportion of trials in which neural response to the unmodulated noise carrier exceeded the criterion level is directly comparable to false alarm rate. This procedure was repeated at 100 criteria values spanning the full range of both distributions. The two-dimensional plot of all pairs of neural hit and false alarm rates forms the neural ROC, and the area under the ROC is called the neural ROC area. Neural ROC area represents neural detectability of a signal—the probability with which an ideal observer can detect a signal (AM) based solely on neural responses. ROC values of 1.0 indicate that the responses to AM were always larger than to unmodulated noise (i.e., the trial-by-trial distribution of firing rate in response to AM noise was higher and had no overlap with the trial-by-trial distribution of firing rate in response to unmodulated noise). ROC values of 0 indicate that the responses to AM noise were always smaller than to unmodulated noise. ROC values of 0.5 indicate that the responses to AM noise and to unmodulated noise were indistinguishable (i.e., the distributions completely overlapped).

Discriminability index, d′

Discriminability index, d', also measures how well a neuron discriminates AM from its unmodulated carriers by comparing distributions of neural response to AM (signal) and unmodulated noise (noise) (Young and Barta, 1986; Middlebrooks and Snyder, 2007; Rosen et al., 2010). Unlike ROC analysis, d' is parametric with an assumption that signal and noise distributions are normal and have the common variance. It is defined as $d' = (\mu_s - \mu_n)/\sigma$, where μ_s and μ_n are mean neural response (firing rate or phase-projected VS) to AM signal and unmodulated noise, respectively, and σ is the common standard deviation (SD). σ was estimated as $\sigma = (\sigma_s + \sigma_n)/2$, where σ_s and σ_n are the SDs of neural response (firing rate or VSpp) to signal and noise, respectively. d' measures the separation of mean values between signal and noise distributions in units of their average SD. Unlike ROC area, d' cannot be converted to probability of detection for direct comparison of unit recordings with behavior. However, d' is not bounded by 1 and 1, and continues to increase proportionally when noise and signal distributions are very well separated. ROC area reaches a ceiling at the value of 0 or 1 (probability correct cannot exceed 1) regardless of how far apart two completely separated distributions are. In contrast, d' has no such bound and will capture the distance between well separated distributions. Thus, in this paper we are not using d' to determine a statistical probability but as an estimate of the overlap of two distributions so it is serving as a measure of separation more than a statistic. We introduced it because of problems when ROC area runs into ceiling effects at values approaching 1.0.

Vector strength and phase-projected vector strength

Vector strength, VS (Goldberg and Brown, 1969), is defined as

$$VS = \frac{\sqrt{\left(\sum_{i=1}^{n} \cos\theta_i\right)^2 + \left(\sum_{i=1}^{n} \sin\theta_i\right)^2}}{n} \qquad (1)$$

A Firing rate-based neural thresholds

B VSpp-based neural thresholds
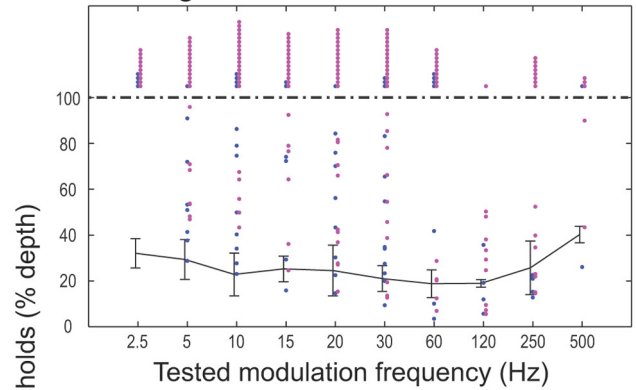
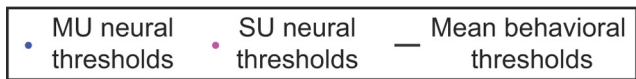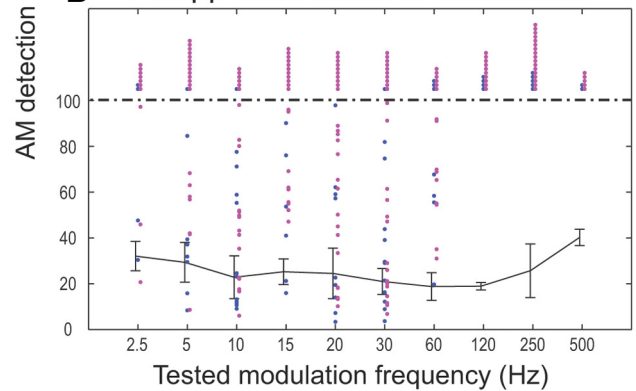| MU neural thresholds | SU neural thresholds | — Mean behavioral thresholds |

Figure 1. Average behavioral thresholds plotted against distribution of single-unit (SU) and multiple-unit (MU) thresholds. **A**, Firing rate thresholds for each SU (pink dot) and MU (blue dot) are superimposed on a plot of the average behavioral thresholds as a function of modulation frequency. Each behavioral point (threshold) is the average of all sessions in which units were recorded regardless of animal. Error bars are SD by session. Points above the 100% mark are units that did not reach the threshold criterion of ROC area >0.75. **B**, Same as **A**, but for phase-projected vector strength. Note no data points are shown for 1000 Hz because no units had 1000 Hz BMF.

where n is the total number of spikes, and θ_i is the phase of each spike in radians.

θ_i is calculated by

$$\theta_i = 2\pi \frac{(t_i \; modulo \; p)}{p} \qquad (2)$$

where t_i is the time of the spike relative to the onset of the stimulus and p is the modulation period of the stimulus. VS measures how tightly the response is temporally locked to one phase of modulation. If all spikes fire at precisely the same phase relative to the stimulus AM, VS is 1. If spikes are circularly symmetric with respect to stimulus phase (this includes random timing), VS will be 0. One weakness of VS is that it may give spuriously high values at low firing rates. If a cell fires one spike on a given trial, a VS value of 1 would result. If a cell fires two spikes randomly, a high VS would also likely result because the probability of two random spikes firing 180° out of phase with each other (relative to the stimulus modulation period) is low. Basically, if sampling from a random distribution of spikes in time, VS will approach zero as the number of spikes approach infinity. Since we apply VS on trial-by-trial basis, VS in low spike-count trials is a critical issue because some single units fire only a few spikes in a single trial.

Neurometrics have been derived that can be used to look at whether temporal patterns can detect or discriminate sounds (Walker et al.,

2008). In this paper, we wanted to focus on the specific temporal code of the ability of neurons to temporally follow the stimulus envelope (phase locking). One way to address this issue is to use a measure called phase-projected vector strength (VS_{PP}) (Yin et al., 2011), which allows for quantification of phase locking on trial-by-trial basis. Unlike VS, VS_{PP} does not give spurious values on low spike trials. Obtaining single-trial measurements is essential for neurometric analysis, where trial-by-trial variance is a key element to compare neural to behavioral results. Determining the best measure to quantify temporal following can be difficult (Kajikawa and Hackett, 2005; Malone et al., 2007, 2010), but here we wanted to examine whether task engagement can improve a neuron's ability to follow temporal modulation of a sound. Conceptually, VS_{PP} compares the mean phase angle for each trial with the mean phase angle of all trials (at 100% depth for corresponding condition) and penalizes single-trial VS values if they are not in phase with the global response. VS_{PP} was calculated on a trial-by-trial basis as follows:

$$VS_{pp} = VS_t \cos(\varphi_t - \varphi_c) \qquad (3)$$

where VS_{PP} is the phase-projected vector strength per trial, VS_t is the vector strength per trial, calculated as in Equation 1, and φ_t and φ_c are the trial-by-trial and mean phase angles in radians. Phase angles φ are calculated as

$$\varphi = arctan2 \frac{\sum_{i=1}^{n} \sin \theta_i}{\sum_{i=1}^{n} \cos \theta_i} \qquad (4)$$

where n is the number of spikes per trial (for φ_t) or across all trials (for φ_c) and arctan2 is a modified version of the arctangent that determines the correct quadrant of the output based on the signs of the sine and cosine inputs (MatLab, atan2; The MathWorks). The mean phase angle φ_c for each cell was estimated from its response to 100% AM. For all VS_{PP} calculations, a cell that fired no spikes was assigned a VS_{PP} of zero. Whereas VS ranges from 1 (all spikes occur at the same phase with respect to stimulus) to 0 (spikes timed randomly with respect to stimulus phase, or spikes occurring circularly symmetric with regard to stimulus phase), VS_{PP} may range from 1 (all spikes in phase with the mean phase) to −1 (all spikes 180° out of phase with mean phase), with 0 corresponding to random phase with regard to the mean phase. Except for the cases in which there were low spike counts, the two VS measures were in good agreement (Yin et al., 2011).

Calculation of neural and behavioral thresholds. On each recording day we used all behavioral trials to calculate hit rates at each depth and false alarm rates. Then we estimated ROC area (Green and Swets, 1974) by calculating the trapezoidal area under the false alarm versus the hit rate curve at each depth (O'Connor et al., 2000). We then plotted ROC area versus depth and fit a sigmoid function to the data points. Threshold was defined as the point at which the sigmoid fit crossed an ROC area of 0.75.

Neural thresholds were calculated using the neural ROC area. For each unit, we calculated a depth sensitivity (neural ROC area vs depth) function for both firing rate (over the 800 ms stimulus) and phase-projected vector strength (VS_{PP}) comparing the modulated stimulus responses to the unmodulated stimulus responses (Johnson et al., 2012). ROC area
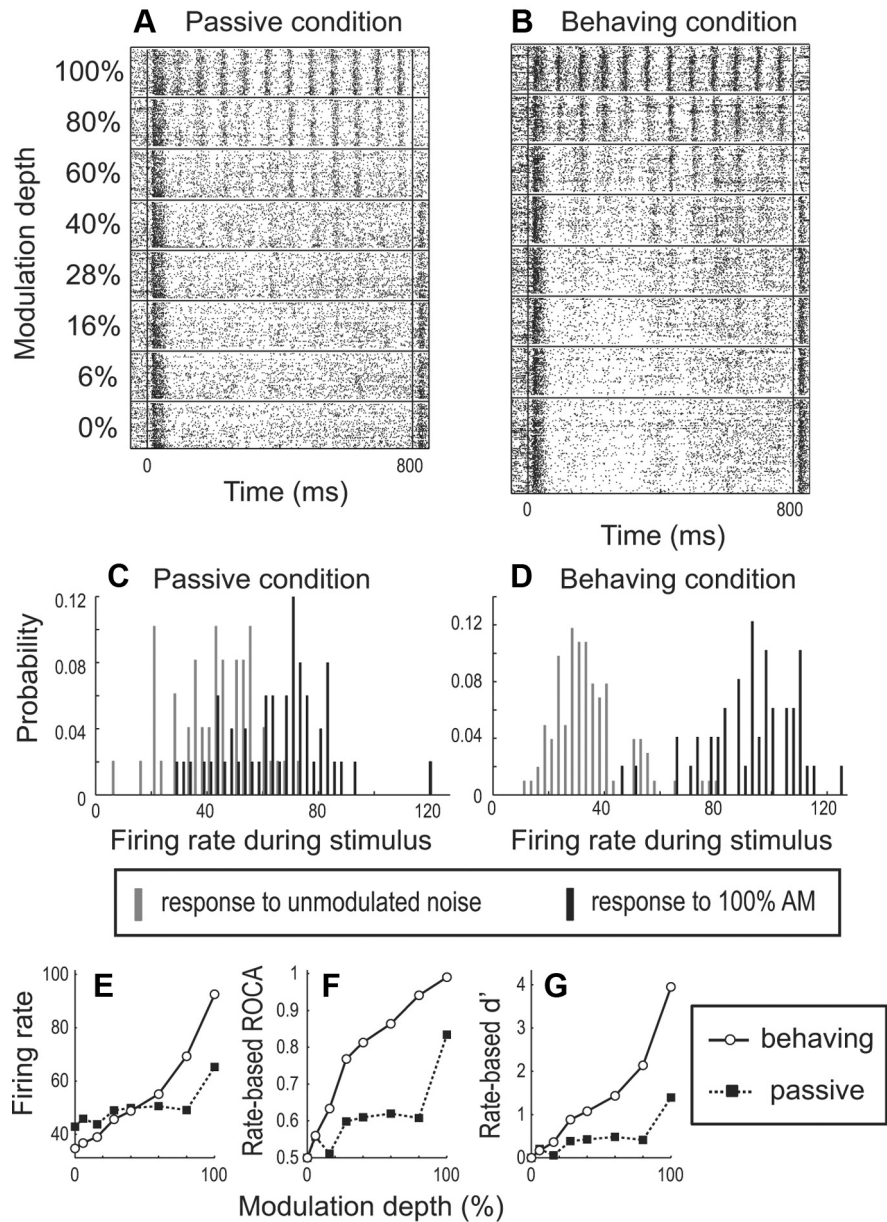

Figure 2. *A*, *B*, Raster plots of a MU response to 15 Hz AM as modulation depth is varied from 100% (top) to 0% (bottom) in the passive (*A*) and the behaving (*B*) conditions. *C*, *D*, Probability distributions of single-trial firing rate (0 – 800 ms) in response to 100% (black bars) and 0% (gray bars) AM in the passive (*C*) and the behaving (*D*) conditions. *E*, Average firing rate during the stimulus is plotted as a function of modulation depth for the passive (solid square) and behaving (open circle) conditions. *F*, *G*, ROC area (*F*) and *d′* (*G*) based on firing rate are plotted as a function of modulation depth for passive (solid square) and behaving (open circle) conditions.

was plotted as a function of modulation depth for each unit. We then fit a sigmoid to the data points and calculated thresholds from these functions as we did for behavior (where the fit crosses an ROC area of 0.75).

Categorization of synchronized and nonsynchronized responses. A synchronized response was defined as one that significantly phase locked at any modulation depth (6–100%). This was quantified by comparing the VS_{PP} values in response to the AM stimulus to the VS_{PP} values for the unmodulated noise response. This was done at each depth with a *t* test and a correction for multiple comparisons (7 depths, *p* corrected to 0.05 by using 0.0073 for each comparison).

A nonsynchronizing response was defined as a response that did not meet the above synchronizing requirement and that could distinguish AM from unmodulated noise. This was objectively quantified as having a significantly different firing rate (measured over the entire 800 ms stimulus duration) in response to AM at any modulation depth (6–100%)
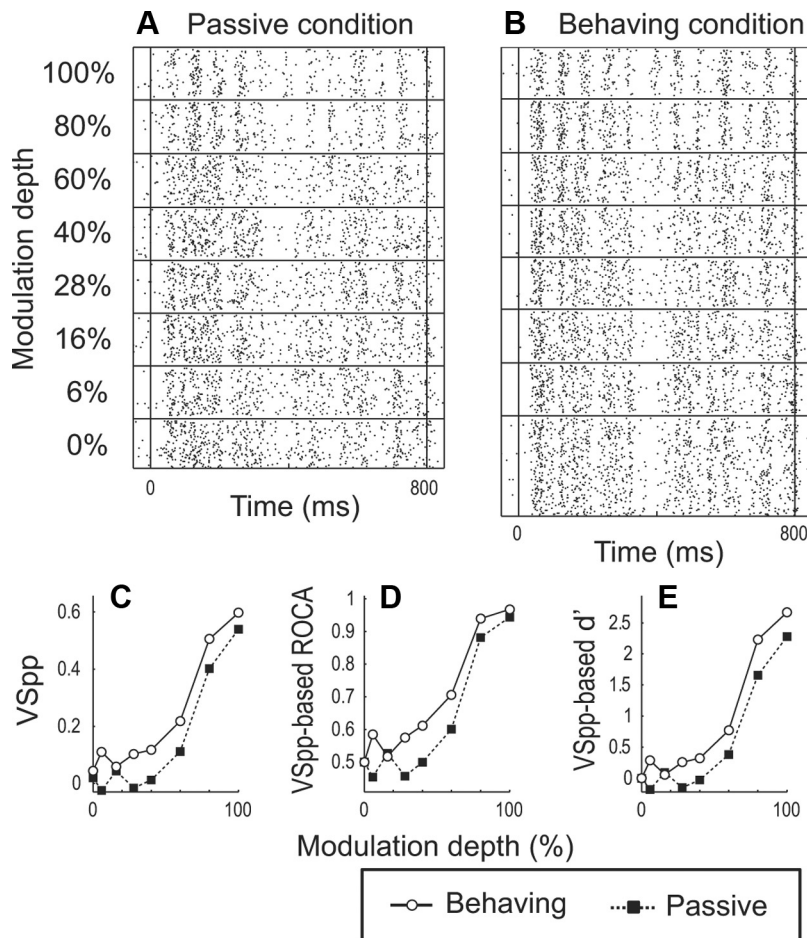
Figure 3. **A**, **B**, Raster plots of an SU response to 15 Hz AM in the passive (**A**) and the behaving (**B**) conditions. **C**, Phase-projected vector strength (VSpp) is plotted as a function of modulation depth for the passive (solid square) and the behaving (open circle) conditions. **D**, **E**, ROC area (**D**) and d' (**E**) based on VSpp are plotted as a function of modulation depth.

intensity combination) plus 2 SD (MatLab, contourc function; The MathWorks). The best frequency (BF) and threshold were determined from the obtained frequency tuning curve. A tonotopic map was created from BF in all recordings for each animal. The location of A1 was determined based on a systematic increase in BF from anterior to posterior axis.

On termination of the experiments, three locations were marked in one animal (monkey V) by inserting electrodes dipped in biotinylated dextran amine. These three locations were at the anterior, middle, and posterior parts on the physiologically determined border between A1 and the middle-medial belt cortex. Then the monkey was given an overdose of sodium pentobarbital and was perfused with 4% paraformaldehyde in 0.1 M phosphate buffer. The brain was removed, blocked, and allowed to sink in 30% sucrose in 0.1 M phosphate buffer before it was frozen. Sections of 50 μm thickness were cut on a sliding microtome in the frontal plane and were alternately processed with three staining methods: treatment with mouse anti-parvalbumin antibody and then with biotinylated horse anti-mouse secondary antibody followed by reactions with acetyl-avidin biotinylated peroxidase complex (ABC) and diamino benzidine (DAB); Nissl staining; and Nissl staining followed by reactions with ABC and DAB. The anatomical boundary of A1 in monkey V was consistent with the physiologically determined borders (O'Connor et al., 2010).

Results

Comparison of behavior to physiology

The animal's behavioral thresholds were in the 15–25% range at the more sensitive modulation frequencies. The behavioral data for unit recording sessions in this study are summarized in Figure 1. Psychometric data for response times (the time from the end of the test stimulus to lever release) and response probability were also recently published (Niwa et al., 2012). The most sensitive MU and SU thresholds, derived using either firing rate or VS-based metrics, were much better than behavioral performance. In Figure 1, unit thresholds are only taken from the behaving condition. The easiest to observe difference between SUs and MUs is the tendency for a higher proportion of SUs not to reach the threshold criterion of ROC area >0.75. It is difficult to compare thresholds between single and multiple units because there is no one best way to handle units that did not reach threshold (Johnson et al., 2012).

than in response to the unmodulated noise (corrected $p < 0.05$). The neural responses from the behaving condition were used for the purpose of these categorizations.

Characterization of A1 and histology. The determination that our recordings were made in A1 was based on the stereotypical tonotopic gradient, the robustness of responses, and the sharpness of frequency tuning (Merzenich and Brugge, 1973; Morel et al., 1993; Kosaki et al., 1997; Rauschecker, 1997; Recanzone, 2000; Recanzone et al., 2000; Kusmierek and Rauschecker, 2009; Yin et al., 2011) obtained from physiological recordings. We also performed histological experiments on one monkey (V) to confirm that our recording sites were located in A1. Two other monkeys (X and W) are still serving as subjects in related experiments and thus are not available for the histological confirmation (Niwa et al., 2012).

Frequency tuning was measured at each recording site by presenting pure tones with different combinations of frequencies and intensities. An initial assessment was made by manually varying frequency and intensity to determine the frequency range used in the automated procedure. For the automated procedure, frequencies typically spanned three octaves with ⅕ octave increments around a center frequency that was estimated by the initial manual assessment. Intensities typically spanned 80 dB with a 10 dB increment between 10 and 90 dB SPL. Tone duration was 100 ms. Stimuli were presented in a random order and repeated at least three times for each frequency–intensity combination. A two-dimensional response matrix (intensity × frequency) was obtained using firing rate during the 100 ms stimulus window. The unit's frequency tuning curve was estimated using the contour line at the mean spontaneous response (spike count in a 75 ms window before the onset of each frequency–

Neural sensitivity to AM improves due to behavioral engagement in the AM task

The ability of an observer to decode neural responses to determine whether a sound was amplitude modulated based on firing rate significantly improved when animals performed an AM discrimination task (behaving condition) compared to when they were passively presented with the same stimuli (passive condition). The task was to determine whether a sound (0–100% depth sinusoidally amplitude modulated noise) was modulated or unmodulated (this task is also referred to as modulation detection). Many units improved their ability to discriminate AM from its unmodulated noise carrier in the behaving condition (Fig. 2).

This MU's responses temporally followed (i.e., phase locked to) the 15 Hz modulation frequency at higher modulation depths in both passive (Fig. 2A) and behaving (Fig. 2B) conditions. The average firing rate also increased monotonically with modulation depth in both conditions. Although the basic response properties appear similar, this unit's AM responses were not the same between conditions. For example, in the behaving condition, the average firing rate (0–800 ms) to 100% AM was higher and to 0% (unmodulated) was lower than in the passive condition (Fig. 2E).

To determine whether the change in mean firing rate between behaving and passive conditions translates to improvement of neuronal AM discrimination performance, we used ROC analysis to calculate neuronal ability to discriminate AM (6–100%) from unmodulated sounds (0%). Average firing rate over the entire stimulus was calculated in each trial, and probability distributions of firing rates were created for unmodulated noise and AM at each depth (Fig, 2C,D). In this example, the distributions of responses to 100% AM (black bars) and unmodulated noise (gray bars) were well separated in the behaving condition but less so in the passive condition, indicating that this MU's firing rate discriminated 100% AM from noise better in the behaving than in the passive condition. ROC area was calculated by comparing distributions of AM responses at each modulation depth to unmodulated noise responses in each condition. Neural ROC area represents the probability that an ideal observer detects modulation based solely on neural responses (ROC of 0.5 = chance). The ROC area for 100% AM was 0.99 in the behaving condition and 0.83 for the passive condition in this example. Rate-based ROC areas were greater in the behaving than in the passive condition at 16–100% depths (Fig. 2F), indicating that an ideal observer using the firing rate of this MU can better discriminate AM from unmodulated noise in the behaving condition. Neural d', an alternative measure of neural discriminability, shows the same effect (Fig. 2G) (the need for using d' is described later).

We also examined whether AM sensitivity based on a temporal code improves when animals engage in the AM detection task. Phase-projected vector strength (VS$_{PP}$; see Materials and Methods for details) was used as a measure of phase locking. VS$_{PP}$ measures the ability of neurons to temporally follow the AM on a trial-by-trial basis (Yin et al., 2011). We did not use standard VS (Goldberg and Brown, 1969) without phase projection because it gives spurious values on low spike trials. Obtaining single-trial measurements is essential for neurometric analysis in which trial-by-trial variance is a key element to compare neural to behavioral results. VS$_{PP}$ was calculated for each trial in the time window excluding onset response (80–800 ms), and VS$_{PP}$-based ROC areas were calculated for the AM (6–100%) in the behaving and passive conditions. An example of an SU that improved VS$_{PP}$-based AM sensitivity is shown in Fig-
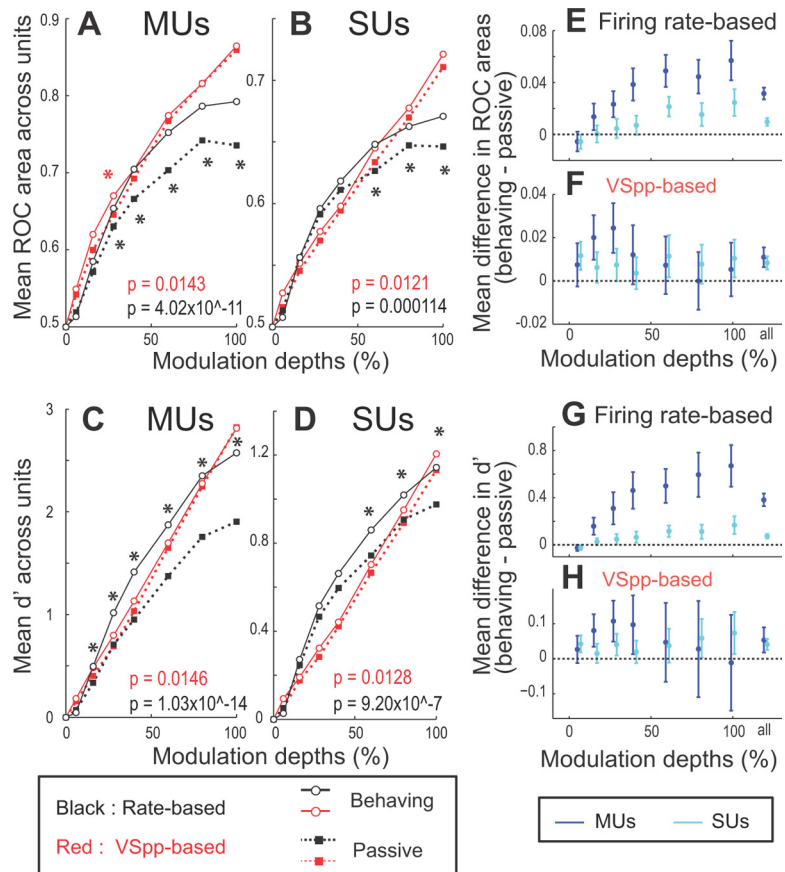

Figure 4. *A, B,* Population-mean ROC areas based on firing rate (black) and VSpp (red) are plotted as a function of modulation depth for the passive (solid square) and the behaving (open circle) conditions for all recorded MUs (*A*) and SUs (*B*). *C, D,* Population mean d' shown for MUs (*C*) and SUs (*D*) in the same manner as in *A* and *B*. In all panels, *p* values are denoted in corresponding colors (red VSpp, black firing rate) showing Wilcoxon signed-rank test comparing across-depth ROC areas between behaving and passive conditions. **p* < 0.05 for individual modulation depths where a Wilcoxon signed-rank test yielded a value. (red asterisk) VSpp. (black asterisk) firing rate, respectively. *E, F,* Plots showing variation in mean difference of ROC area. The difference in ROC area between behaving and passive conditions was calculated for every unit to create a distribution of differences. The mean and standard error are then plotted for each modulation depth, and for all responses collapsed across modulation depth (far right point). *G, H,* Same as *E, F* but this time for d'.

ure 3. This neuron phase locks to the 15 Hz AM at higher modulation depths and monotonically decreases firing rate with increasing depths in both passive (Fig. 3A) and behaving (Fig. 3B) conditions. Trial-averaged VS$_{PP}$, ROC area, and d' all increased in the behaving condition (Fig. 3C–E).

The population average ROC area versus depth functions show that the aggregate of activity across the population of A1 neurons better discriminates AM from unmodulated noise in the behaving than in the passive condition, based either on rate or phase-locking codes (Fig. 4). To test the overall effect of the behavioral condition on ROC area, ROC area data were collapsed across all depths in each condition, and a single Wilcoxon signed-rank test was performed for each condition with each measure (firing rate or VSpp). For MUs, both rate- and VS$_{PP}$-based ROC areas across depths significantly increased in the behaving compared to the passive condition (Fig. 4A–D; p values shown at the bottom of each plot). Mean difference for data collapsed across all depths is shown in the far right column labeled "all" in Fig. 4E,F). For SUs, similar results were obtained (Fig. 4B,E,F). Essentially identical results were obtained with d' (Fig. 4C,D,G,H). The results in Figure 4 indicate that AM sensitivity based on a temporal code, measured by VS$_{PP}$, significantly improves due to the engagement in the AM task; however, the magnitude of improvement is considerably smaller than that for rate-based AM
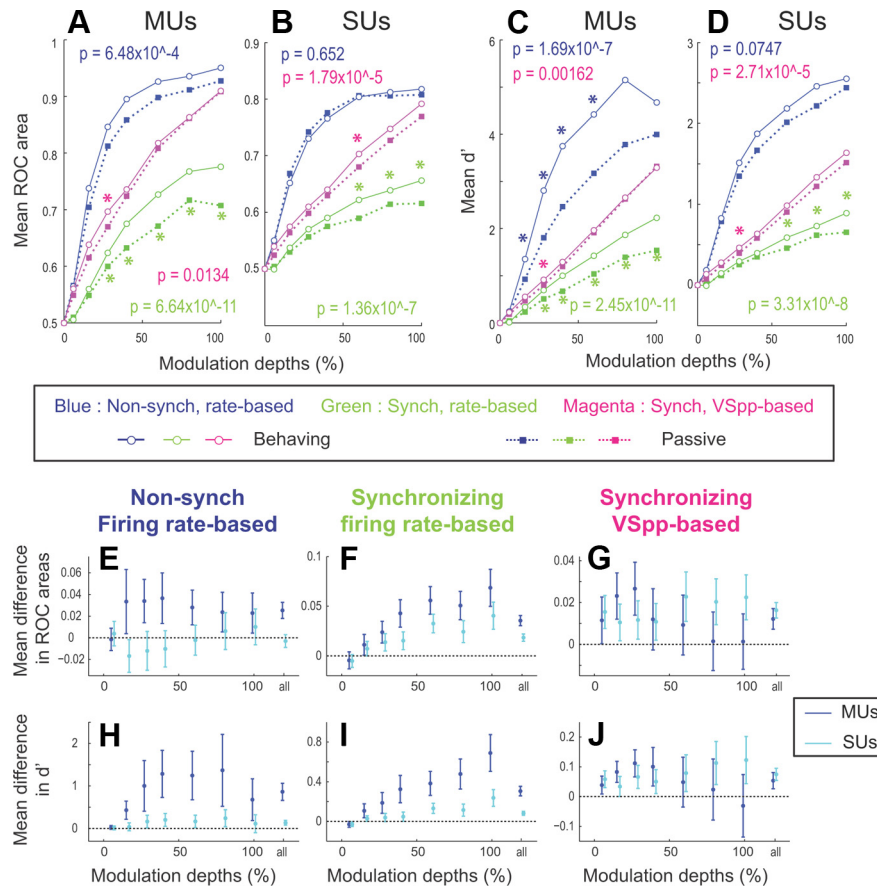
Figure 5. Same style graph as in the previous figure, except in this figure nonsynchronized and synchronized (for rate and VSpp) responses are separately analyzed.

sensitivity. The improved neural ROC areas in the behaving condition suggest that engagement in discrimination does not simply increase firing equally (in the sense of adding a constant to the firing rate) for all depths (0–100%) but, rather, serves to create larger differences in firing rate and phase locking between modulated sounds and the unmodulated sound in the behaving condition.

Engagement in the AM task improves rate-based AM sensitivity of both synchronizing and nonsynchronizing responses

Auditory neurons can encode AM with changes in firing rate and/or synchronizing to the AM temporal envelope (phase-locking). Some neurons respond to AM at some modulation frequencies by changing firing rate without synchronizing action potential timing to the modulation (nonsynchronized responses). Such nonsynchronized responses have been proposed to have a special role in AM perception (Lu et al., 2001; Liang et al., 2002; Bartlett and Wang, 2007; Bendor and Wang, 2007). AM coding by phase locking (synchronizing responses) often accompanies a change in firing rate. Thus, synchronizing responses may use both rate and phase-locked codes while nonsynchronizing responses are thought to use rate to encode AM. Here, we examined whether AM sensitivity of synchronizing and nonsynchronizing responses was differentially affected by the change in behavioral states.

AM sensitivity significantly improves for synchronizing responses in A1 (Fig. 5). The phase-locked based ROC area significantly improved for behaving relative to passive conditions across depth for both SUs and MUs and for both ROC area and d'

(Fig. 5 A,B, magenta, G). This is different from responses in Figure 4 in which all units (synchronized, nonsynchronized, and those that did not show significant changes in firing rate or phase locking with modulation depth) are included, whereas the magenta and green curves in Figure 5A–D,G,J only include synchronized responses. The rate-based ROC areas for synchronizing responses are significantly greater in the behaving condition compared to the passive condition across depths (Fig. 5A,B, green, F). For nonsynchronizing responses, there was a significant increase in across-depth, rate-based ROC areas of MUs, but not of SUs (Fig. 5A,B top, blue, E). Note that for the population average, rate-based ROC areas are much higher for nonsynchronized responses than those for synchronized responses. This means that many of the nonsynchronized responses in the passive state can already distinguish AM from unmodulated noise very well at higher depths. Since ROC areas are bounded by the values of 0 and 1, there may be a ceiling effect on the change in AM sensitivity for cells that have well separated distributions of AM and unmodulated noise in the passive condition. To investigate this possibility, we used d' as an alternate statistic because when distributions of unmodulated and modulated responses are completely separated, ROC area gives values of 0 or 1 regardless of how far apart those distributions are, while d' is unbounded and gives values reflecting the distance between those distributions. When nonsynchronized responses are analyzed with d', the MUs showed a robust significant change, but the SU improvement only approached significance, showing a trend of increase in the behaving condition for across-depth data ($p = 0.0747$ by Wilcoxon sign-rank test) (Fig. 5D, blue lines, H).

Together, the results indicate that rate-based AM discriminability improves due to the engagement in the AM task for both synchronizing and nonsynchronizing responses. Although the nonsynchronizing responses of SUs did not show a significant increase in d' or in ROC area, MUs did show a significant increase. In addition, nonsynchronizing responses have much better AM sensitivity compared to synchronizing responses regardless of behavioral states, implicating the importance of nonsynchronizing responses in AM detection.

Improvement in rate-based AM sensitivity is greater in the earlier period of stimulus presentation for suprathreshold stimuli

We conducted ROC analyses using different time windows and found that the improvement in AM sensitivity due to behavioral states is time dependent. Rate-based ROC areas were calculated in time windows of 400 ms duration beginning at 0, 100, 200, 300, and 400 ms after stimulus onset. We found that for 100 and 80% AM stimuli, differences in rate-based ROC areas between the behaving and passive conditions appear greater in the first half of the stimulus (0–400 ms) than in the second half (400–800 ms), whereas this effect does not appear to be present at lower modu-

lation depths (Fig. 6). Two-way ANOVA examining the effect of time (first or second half of stimulus) and condition (behaving or passive) confirmed this for MUs (Table 1; significant interaction at 100 and 80%). For 40, 28, and 16%, the difference of ROC areas between conditions (behaving or passive) stayed approximately constant for different starting times of the analysis windows (Table 1; those interactions were not significant). This result suggests that the modulation of AM sensitivity by the behavioral state in A1 neurons may have the biggest impact during the earlier period of stimulus presentation for stimuli with larger depths (i.e., presumably more easily detectable AM). Although there are many potential explanations, an intriguing possibility is that this might be because the discrimination could be made earlier during the stimulus for these sounds, and the effect of engagement in the AM detection task might be reduced in the later period. In contrast, those stimuli near the animals' behavioral AM detection thresholds (~20–25% depth) show constant improvement in AM sensitivity over the stimulus period, possibly because the detection of modulation takes longer and, therefore, requires a longer period of engagement.

Figure 6. *A*, *B*, Rate-based ROC areas are calculated in 400 ms time windows starting at various times after stimulus onset. Population mean rate-based ROC area is plotted as a function of start time of the time windows at each modulation depth (from 100% on the left to 6% on the right) for MUs (*A*) and SUs (*B*).

Firing rate relative to spontaneous activity

Thus far, we have shown how the ability to detect modulation as measured by ROC analysis improves with task engagement. A related question is whether and how raw firing rates change. We found that both the spontaneous and the driven firing rates to AM were higher in the behaving compared to the passive condition (Fig. 7*A*,*B*), although the firing rate in response to the unmodulated (0% depth) sound was not significantly different between the passive and the behaving conditions. The differences between behaving and passive conditions appear to get larger with increasing modulation depth. Interestingly, when the appropriate spontaneous rates were subtracted from the driven firing rate, the spontaneous-adjusted rate was lower in the behaving than the passive condition (Fig. 7*C*,*D*). However, the slope of the spontaneous-adjusted firing rate versus the depth function was steeper in the behaving condition compared to the passive condition (Fig. 7*C*,*D*), which supports the improved ability of neurons to distinguish between modulated and unmodulated sounds during task engagement.

Comparison of MUs and SUs

While Figure 1 shows that the most sensitive MUs and SUs have similar thresholds, Figures 4–6 paint a slightly different picture. In Figures 4–6, mean *d′* and ROC areas are higher for MUs than SUs. Unlike the case for thresholds (Fig. 1) the mean *d′* and ROC areas for the most AM-responsive SUs and MUs were not similar; MUs had higher *d′* values and ROC areas (Fig. 8). Here most AM-responsive SUs (or MUs) are defined by averaging *d′* or ROC area across all depths and both conditions and taking those with the highest 25% values for this average. Although we cannot be certain, one explanation for MU improved ability to discriminate AM from unmodulated noise could be the added statistical reliability gained by pooling across the multi-unit. A pooling effect
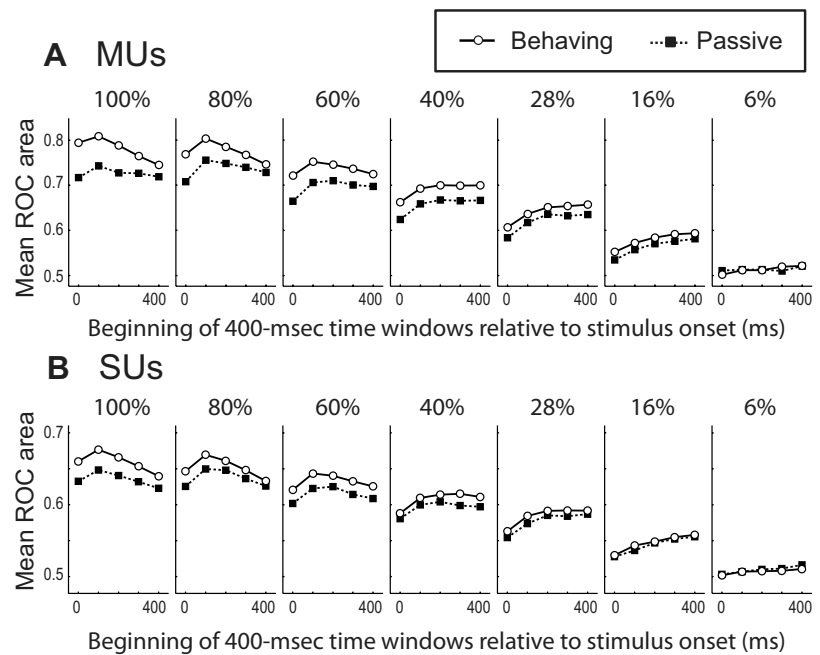
that improved neural sensitivity using similar stimuli in passive animals has recently been demonstrated (Johnson et al., 2012).

In addition to their increased ability to distinguish AM from unmodulated noise, MUs show greater differences between *d′* values (also ROC areas) for firing rate in the behaving and passive conditions than do SUs (Fig. 4, compare *C*, *D*; also see Fig. 4*E*,*G* and 5*E*,*F*,*H*,*I*). One salient difference between SUs and MUs is the higher proportion of SUs that do not reach threshold (Fig. 1). It is possible that during averaging the larger number of nonsensitive SUs reduces the average difference between behaving and passive conditions for SUs, especially if these nonsensitive SUs do not distinguish between the two conditions (we will call this "smearing"). This might then lead to the observed smaller differences for SUs in Figs. 4*E*,*G*, 5*E–I*. It also might be that the SUs with the highest mean *d′* and ROC areas are as good as the best MUs. To investigate these possibilities, we calculated mean ROC areas and mean *d′* values for the most AM-responsive SUs (shown in Fig. 8). If nonsensitive SUs are responsible for a reduction in the behaving–passive difference, the most AM-responsive SUs (top 25%) should show a greater behaving–passive difference than the overall population, but this is not the case. For *d′*, the behaving–passive difference is similar for both the top 25% of cells and the overall population (Fig. 8*G*,*H*). For ROC area, the top 25% of SUs, if anything, have higher values in the passive than the behaving condition (Fig. 8*D*). These results argue against the smearing case and also against the best SUs contributing more to the differences observed between behaving and passive conditions. Interestingly, these results, particularly the ROC area results, suggest that the most AM-responsive units may in fact reduce overall mean differences between behaving and passive conditions.

The fact that for ROC area there is little difference between behaving and passive conditions could reflect ceiling effects. As a measure, ROC area is bounded at 1, so increases in sensitivity for already sensitive units may be limited; *d′*, on the other hand, is

Table 1. *p* values of two-way repeated measures ANOVA on behavioral condition and time

	100%	80%	60%	40%	28%	16%	6%
MUs							
Behaving versus passive	0.0011*	0.0023*	0.0004*	0.00018*	0.0142*	0.1307	0.3628
1st versus 2nd half	0.0523	0.9024	0.1791	0.0045*	0.0005*	0.0005*	0.0350*
Interaction	0.0004*	0.0209*	0.0543	0.5085	0.7039	0.6758	0.3647
SUs							
Behaving versus passive	0.0489*	0.2000	0.0202*	0.2169	0.6296	0.8370	0.2057
1st versus 2nd half	0.0588	0.3218	0.6509	0.0432*	0.0003*	0.0000*	0.0219*
Interaction	0.1741	0.3023	0.9701	0.4025	0.5989	0.7383	0.7059

We obtained *p* values from two-way ANOVA (repeated measures) testing for the effects of behavioral conditions (factor 1; behaving vs passive) and time window conditions (factor 2; first vs second halves of the test stimulus) on the firing rate-based ROC areas. We also determined the interaction of these factors (factor 1 × factor 2). *Significant at $p < 0.05$.
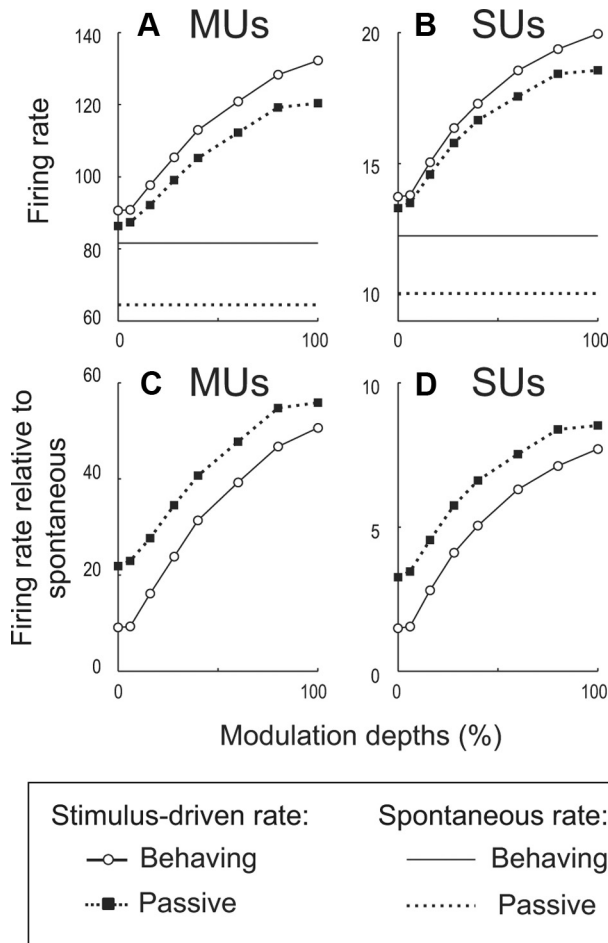

Figure 7. *A*, *B*, Population mean stimulus-evoked firing rate is plotted as a function of modulation depth for the passive (solid squares) and behaving (open circles) conditions for all recorded MUs (*A*) and SUs (*B*). Dotted and solid horizontal lines denote spontaneous firing rate in the passive and behaving conditions, respectively. *C*, *D*, Stimulus-evoked firing rate relative to spontaneous firing rate (firing rate spontaneous, where spontaneous is different for behaving and passive conditions) is plotted as a function of modulation depth.

not bounded, and increases in sensitivity can be seen regardless of the initial sensitivity of the unit.

Discussion

These results show that at the SU and MU levels, task engagement (1) not only increases activity but changes the ability to tell two stimuli apart and (2) does not act solely on average firing rate but also can improve temporal codes by increasing the temporal precision of firing.

Effect of auditory task engagement on auditory cortical response

Auditory cortical stimulus-evoked and spontaneous firing rates depend on behavioral state (Miller et al., 1972; Hocherman et al., 1976; Pfingst et al., 1977; Benson and Hienz, 1978; Benson et al., 1981; Scott et al., 2007; Otazu et al., 2009; Jaramillo and Zador, 2011). Changes in driven and spontaneous firing rates between behaving and passive conditions are inconsistent across studies (varying from suppression, to no change, to enhancement). The discrepancy may reflect differences in behavioral task (for review, see Sutter and Shamma, 2011) as it is well known that small differences in task and stimulus configuration can have large effects on modulation of neural activity (e.g., Groh et al., 1996; Boudreau et al., 2006). We found both evoked and spontaneous firing rates were raised in the behaving compared to the passive condition (Fig. 7*A*,*B*). Otazu et al. (2009) proposed that tasks tend to have selective and nonselective attentional demands that increase and decrease activity, respectively, and our physiological results are consistent with this model if the AM discrimination used engaged selective attention.

Active engagement has also been shown to change A1 neuron tuning properties compared to passive conditions (Fritz et al., 2003; Atiani et al., 2009; Lee and Middlebrooks, 2011). Lee and Middlebrooks (2011) found that A1 neuron spatial tuning becomes sharper, likely due to the suppression of responses to less preferred locations, when animals localize sound. Fritz et al. (2003) found many A1 neurons exhibit facilitative changes in their spectro-temporal receptive fields (STRFs) when animals performed tone detection. The facilitative change took the form of an increase in excitation or a reduction in inhibitory sideband near the frequency of target tones in the detection task.

Using neurometric analysis we found that neurons can better distinguish modulated from unmodulated sounds during discrimination (behavior) and that the improved neural discriminability is not due to a general increase in firing rate during the behaving condition but rather to stimulus-dependent changes. Firing rate increased more for modulated than unmodulated sounds in the behaving condition, rendering the two more distinguishable.

Because both evoked and spontaneous firing rates are higher in the behaving compared to the passive condition, an interesting relationship is observed. Spontaneous-adjusted driven rate is *lower* in the behaving condition (Fig. 7*C*,*D*). Then why does neural discriminability increase? It increases because during task performance the neurons barely respond above spontaneous to unmodulated noise but respond much more to modulated sounds; therefore, modulation contrast is improved. This result is consistent with that of Atiani et al. (2009), who focused on spectral, more than temporal, contrast. They trained ferrets to detect a tone within a spectrally complex sound. Following training they found differences in the STRF between active and passive condi-

tions. The changes had two components: a decrease in overall activity relative to spontaneous (gain shift) and a frequency selective increase in driven activity at frequencies near the target for high signal-to-noise trials.

Additionally, we observed that AM discriminability based on phase-locking improves in the behaving condition. To our knowledge, this is the first demonstration of an improvement in an auditory neuron's ability to follow stimulus temporal structure when performing a task that requires attention to this structure.

Relationship to sound processing

Our data can be interpreted in the context of two problems encountered by the auditory system. One is, can more than one sound feature be simultaneously encoded? The other is, what is the best way to encode AM? Under more natural conditions these can be combined into, how does the brain best represent one sound feature in a complex stream of many others?

Relationship to sound processing: coding multiple sound features

A1 has the challenge of simultaneously encoding multiple sound features to be sent to more specialized parallel pathways higher in the auditory system (Rauschecker and Tian, 2000; Woods et al., 2006; Leaver and Rauschecker, 2010; Hackett, 2011). This can be achieved in several ways. If neurons use only firing rate, a population code is required because a single neuron's firing rate cannot unambiguously code several features simultaneously. For example, for neurons that respond to AM by monotonically increasing firing rate for modulation depth and loudness, intermediate firing rates could either indicate loud low-depth or soft high-depth sounds. One population code is that multiple features simultaneously present in a sound can be encoded by distinct "feature-detecting" neurons. The features can be recovered by observing which neurons fire (Barlow, 1972; Suga, 1989; Groh, 2001). An alternative to this one-neuron one-feature scheme is that each neuron encodes multiple parameters, but the population is needed to disambiguate the fact that a single firing rate cannot uniquely identify the parameter (Rolls et al., 1997; Petkov et al., 2007; Bizley et al., 2009).

Two results suggest a special role for nonsynchronized responses as AM "feature detectors" working in parallel with other neurons to represent multi-parameter sounds (Lu et al., 2001). First, the ability of MU nonsynchronized firing rate to discriminate AM from unmodulated noise significantly improves when the animal is discriminating. This suggests that attention can modulate nonsynchronized responses. Second, nonsynchronized rates are very sensitive to AM (Fig. 5) and, therefore, better at detecting modulation than synchronized responses.

A different method of encoding multiple features simultaneously is temporal multiplexing, by which different features are coded separately within the same neuron, embedded in different time scales,
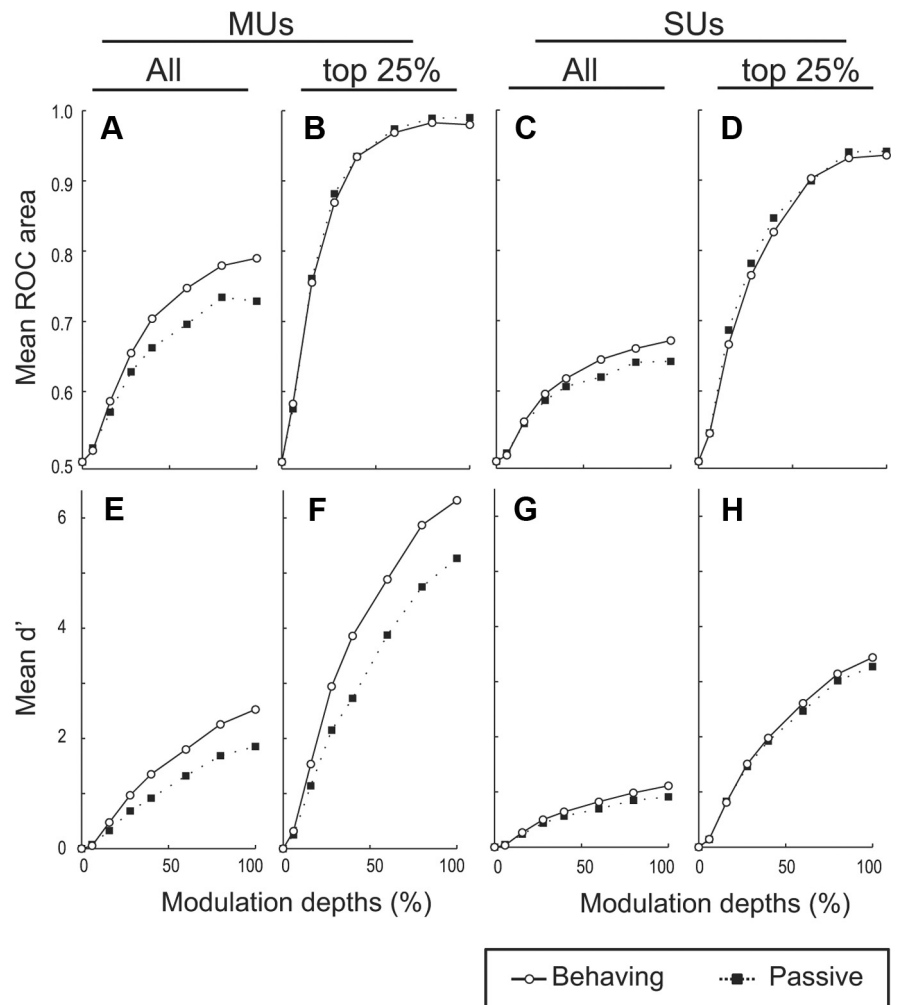

Figure 8. *A–D*, Comparison of mean ROC area for MUs and SUs between the units that best discriminated AM (*B*, MUs; *D*, SUs) and the entire population of units (*A*, MUs; *C*, SUs). The top 25% was determined for each plot by averaging ROC (top), including both conditions and all depths (averaging all responses of the unit regardless of depth and behavioral condition) and selecting the units whose averaged ROC was in the highest quartile. *E–H*, Same as *A–D*, but using d'.

ranging from average firing rate to millisecond-precision spike timing (Sutter and Margoliash, 1994; Gawne et al., 1996; Victor, 2000; Fairhall et al., 2001; Elhilali et al., 2004; Chase and Young, 2006; Ahissar and Knutsen, 2008; Panzeri et al., 2009; Walker et al., 2011). A multiplexed code's advantage over a single time scale code is that a neuron can simultaneously represent different features, thereby increasing a single neuron's coding capacity. Synchronized responses appear to use a multiplexed code by which average firing rate and phase locking can encode different information (Lu et al., 2001; Salinas and Sejnowski, 2001; Friedrich et al., 2004; Yin et al., 2011). The evidence for multiplexed coding in synchronized responses requires more careful interpretation because they can use rate or temporal codes, while nonsynchronized responses are thought to use only rate coding. VS is more sensitive than firing rate in the passive condition, but rate sensitivity improves more with task engagement, bringing sensitivity between the two measures closer (Figs. 4, 5, green vs magenta). This relationship between phase locking and rate is similar to that obtained in primary somatosensory cortex for low flutter frequencies (Salinas et al., 2000). When this is combined with evidence that firing rate is more tightly coupled to behavior and decisions (Hernandez et al., 2000; Niwa et al., 2012), it is reasonable to interpret firing rate as a higher level code extracted from an accurate temporal representation.

Complementary coding of one sound feature

Another way to view the simultaneous use of temporal and rate codes is as complementary codes that together provide more information about one sound feature (Furukawa and Middlebrooks, 2002; Nelken et al., 2005; Atencio et al., 2008; Kayser et al., 2009; Bizley et al., 2010; Shih et al., 2011). As such, higher level neurons could combine rate and temporal information to encode more accurately than either alone. Our results can be interpreted as synchronized responses using complementary rate and temporal codes (both of which can be improved by behavioral state) to represent AM more accurately. This complementary representation could combine at higher levels to create more sensitive, feature selective, nonsynchronized responses. Thus, nonsynchronizing responses may reflect higher level feature selectivity as part of a hierarchy. Both levels of this hierarchical processing can be found in the same area, or even in the same neurons. Yin et al. (2011) have shown that many neurons use nonsynchronized rates at some modulation frequencies and synchronized response properties at others, suggesting that neurons are using different schemes at different modulation frequencies. This suggests that this hierarchical processing is not strictly confined by brain area but for AM could gradually emerge through the auditory neuraxis.

Task-specific considerations

We calibrated all AM stimuli to the same intensity, and animals performed each block at one modulation frequency; only modulation depth varied. Therefore, firing rate could uniquely identify depth. Had a range of intensities, modulation frequencies, and depths been used, a single neuron could not solely rely on firing rate to determine whether the sound was modulated. Then phase-locking, which is less-dependent on mean intensity, might be more important, and greater task-dependent modulation of VS might occur. In addition, multiplexed coding schemes could be directly addressed if multiple stimulus parameters are varied. If multiple codes are in place, it will be helpful to design behavioral experiments that exploit limitations of different codes to manipulate the relevance of the different codes in solving the task.

References

Ahissar E, Knutsen PM (2008) Object localization with whiskers. Biol Cybern 98:449–458.

Atencio CA, Sharpee TO, Schreiner CE (2008) Cooperative nonlinearities in auditory cortical neurons. Neuron 58:956–966.

Atiani S, Elhilali M, David SV, Fritz JB, Shamma SA (2009) Task difficulty and performance induce diverse adaptive patterns in gain and shape of primary auditory cortical receptive fields. Neuron 61:467–480.

Barlow HB (1972) Single units and sensation: a neuron doctrine for perceptual psychology? Perception 1:371–394.

Bartlett EL, Wang X (2007) Neural representations of temporally modulated signals in the auditory thalamus of awake primates. J Neurophysiol 97:1005–1017.

Bendor D, Wang X (2007) Differential neural coding of acoustic flutter within primate auditory cortex. Nat Neurosci 10:763–771.

Benson DA, Hienz RD (1978) Single-unit activity in the auditory cortex of monkeys selectively attending left vs right ear stimuli. Brain Res 159:307–320.

Benson DA, Hienz RD, Goldstein MH Jr (1981) Single-unit activity in the auditory cortex of monkeys actively localizing sound sources: spatial tuning and behavioral dependency. Brain Res 219:249–267.

Bizley JK, Walker KM, Silverman BW, King AJ, Schnupp JW (2009) Interdependent encoding of pitch, timbre, and spatial location in auditory cortex. J Neurosci 29:2064–2075.

Bizley JK, Walker KM, King AJ, Schnupp JW (2010) Neural ensemble codes for stimulus periodicity in auditory cortex. J Neurosci 30:5078–5091.

Boudreau CE, Williford TH, Maunsell JH (2006) Effects of task difficulty and target likelihood in area V4 of macaque monkeys. J Neurophysiol 96:2377–2387.

Bregman AS, Levitan R, Liao C (1990) Fusion of auditory components: effects of the frequency of amplitude modulation. Percept Psychophys 47:68–73.

Britten KH, Shadlen MN, Newsome WT, Movshon JA (1992) The analysis of visual motion: a comparison of neuronal and psychophysical performance. J Neurosci 12:4745–4765.

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

Chase SM, Young ED (2006) Spike-timing codes enhance the representation of multiple simultaneous sound-localization cues in the inferior colliculus. J Neurosci 26:3889–3898.

Eggermont JJ (1995) Representation of a voice onset time continuum in primary auditory cortex of the cat. J Acoust Soc Am 98:911–920.

Elhilali M, Fritz JB, Klein DJ, Simon JZ, Shamma SA (2004) Dynamics of precise spike timing in primary auditory cortex. J Neurosci 24:1159–1172.

Engineer CT, Perez CA, Chen YH, Carraway RS, Reed AC, Shetake JA, Jakkamsetti V, Chang KQ, Kilgard MP (2008) Cortical activity patterns predict speech discrimination ability. Nat Neurosci 11:603–608.

Fairhall AL, Lewen GD, Bialek W, de Ruyter Van Steveninck RR (2001) Efficiency and ambiguity in an adaptive neural code. Nature 412:787–792.

Fitch RH, Tallal P, Brown CP, Galaburda AM, Rosen GD (1994) Induced microgyria and auditory temporal processing in rats: a model for language impairment? Cereb Cortex 4:260–270.

Friedrich RW, Habermann CJ, Laurent G (2004) Multiplexing using synchrony in the zebrafish olfactory bulb. Nat Neurosci 7:862–871.

Fritz J, Shamma S, Elhilali M, Klein D (2003) Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. Nat Neurosci 6:1216–1223.

Furukawa S, Middlebrooks JC (2002) Cortical representation of auditory space: information-bearing features of spike patterns. J Neurophysiol 87:1749–1762.

Gawne TJ, Kjaer TW, Richmond BJ (1996) Latency: another potential code for feature binding in striate cortex. J Neurophysiol 76:1356–1360.

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

Grady CL, Van Meter JW, Maisog JM, Pietrini P, Krasuski J, Rauschecker JP (1997) Attention-related modulation of activity in primary and secondary auditory cortex. Neuroreport 8:2511–2516.

Green DM, Swets JA (1974) Signal detection theory and psychophysics. New York: Krieger.

Griffiths TD, Rees A, Witton C, Cross PM, Shakir RA, Green GG (1997) Spatial and temporal auditory processing deficits following right hemisphere infarction: a psychophysical study. Brain 120(Pt 5):785–794.

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

Groh JM (2001) Converting neural signals from place codes to rate codes. Biol Cybern 85:159–165.

Groh JM, Seidemann E, Newsome WT (1996) Neurophysiology: neural fingerprints of visual attention. Curr Biol 6:1406–1409.

Hackett TA (2011) Information flow in the auditory cortical network. Hear Res 271:133–146.

Heffner HE, Heffner RS (1986) Effect of unilateral and bilateral auditory cortex lesions on the discrimination of vocalizations by Japanese macaques. J Neurophysiol 56:683–701.

Heffner HE, Heffner RS (1989) Unilateral auditory cortex ablation in macaques results in a contralateral hearing loss. J Neurophysiol 62:789–801.

Hernández A, Zainos A, Romo R (2000) Neuronal correlates of sensory discrimination in the somatosensory cortex. Proc Natl Acad Sci U S A 97:6191–6196.

Hocherman S, Benson DA, Goldstein MH Jr, Heffner HE, Hienz RD (1976) Evoked unit activity in auditory cortex of monkeys performing a selective attention task. Brain Res 117:51–68.

Hubel DH, Henson CO, Rupert A, Galambos R (1959) Attention units in the auditory cortex. Science 129:1279–1280.

Imaizumi K, Priebe NJ, Sharpee TO, Cheung SW, Schreiner CE (2010) Encoding of temporal information by timing, rate, and place in cat auditory cortex. PLoS One 5:e11531.

Jaramillo S, Zador AM (2011) The auditory cortex mediates the perceptual effects of acoustic temporal expectation. Nat Neurosci 14:246–251.

Johnson JS, Yin P, O'Connor KN, Sutter ML. (2012) The ability of primary auditory cortical (A1) neurons to detect amplitude modulation with rate

and temporal codes: neurometric analysis. J Neurophysiol. Advance online publication.

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

Kajikawa Y, Hackett TA (2005) Entropy analysis of neuronal spike train synchrony. J Neurosci Methods 149:90–93.

Kajikawa Y, de la Mothe LA, Blumell S, Sterbing-D'Angelo SJ, D'Angelo W, Camalier CR, Hackett TA (2008) Coding of FM sweep trains and twitter calls in area CM of marmoset auditory cortex. Hear Res 239:107–125.

Kayser C, Montemurro MA, Logothetis NK, Panzeri S (2009) Spike-phase coding boosts and stabilizes information carried by spatial and temporal spike patterns. Neuron 61:597–608.

Kosaki H, Hashikawa T, He J, Jones EG (1997) Tonotopic organization of auditory cortical fields delineated by parvalbumin immunoreactivity in macaque monkeys. J Comp Neurol 386:304–316.

Kusmierek P, Rauschecker JP (2009) Functional specialization of medial auditory belt cortex in the alert rhesus monkey. J Neurophysiol 102:1606–1622.

Leaver AM, Rauschecker JP (2010) Cortical representation of natural complex sounds: effects of acoustic features and auditory object category. J Neurosci 30:7604–7612.

Lee CC, Middlebrooks JC (2011) Auditory cortex spatial sensitivity sharpens during task performance. Nat Neurosci 14:108–114.

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.

Lu T, Liang L, Wang X (2001) Temporal and rate representations of time-varying signals in the auditory cortex of awake primates. Nat Neurosci 4:1131–1138.

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

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

Merzenich MM, Brugge JF (1973) Representation of the cochlear partition on the superior temporal plane of the macaque monkey. Brain Res 50:275–296.

Middlebrooks JC, Snyder RL (2007) Auditory prosthesis with a penetrating nerve array. J Assoc Res Otolaryngol 8:258–279.

Miller JM, Sutton D, Pfingst B, Ryan A, Beaton R, Gourevitch G (1972) Single cell activity in the auditory cortex of Rhesus monkeys: behavioral dependency. Science 177:449–451.

Morel A, Garraghty PE, Kaas JH (1993) Tonotopic organization, architectonic fields, and connections of auditory cortex in macaque monkeys. J Comp Neurol 335:437–459.

Nelken I, Chechik G, Mrsic-Flogel TD, King AJ, Schnupp JW (2005) Encoding stimulus information by spike numbers and mean response time in primary auditory cortex. J Comput Neurosci 19:199–221.

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.

O'Connor KN, Barruel P, Sutter ML (2000) Global processing of spectrally complex sounds in macaques (Macaca mullata) and humans. J Comp Physiol 186:903–912.

O'Connor KN, Yin P, Petkov CI, Sutter ML (2010) Complex spectral interactions encoded by auditory cortical neurons: relationship between bandwidth and pattern. Front Syst Neurosci 4:145.

O'Connor KN, Johnson JS, Niwa M, Noriega NC, Marshall EA, Sutter ML (2011) Amplitude modulation detection as a function of modulation frequency and stimulus duration: comparisons between macaques and humans. Hear Res 277:37–43.

Otazu GH, Tai LH, Yang Y, Zador AM (2009) Engaging in an auditory task suppresses responses in auditory cortex. Nat Neurosci 12:646–654.

Panzeri S, Brunel N, Logothetis NK, Kayser C (2010) Sensory neural codes using multiplexed temporal scales. Trends Neurosci 33:111–120.

Petkov CI, O'Connor KN, Sutter ML (2003) Illusory sound perception in macaque monkeys. J Neurosci 23:9155–9161.

Petkov CI, O'Connor KN, Sutter ML (2007) Encoding of illusory continuity in primary auditory cortex. Neuron 54:153–165.

Pfingst BE, O'Connor TA, Miller JM (1977) Response plasticity of neurons in auditory cortex of the rhesus monkey. Exp Brain Res 29:393–404.

Rauschecker JP (1997) Processing of complex sounds in the auditory cortex of cat, monkey, and man. Acta Otolaryngol Suppl 532:34–38.

Rauschecker JP, Tian B (2000) Mechanisms and streams for processing of

"what" and "where" in auditory cortex. Proc Natl Acad Sci U S A 97:11800–11806.

Recanzone GH (2000) Response profiles of auditory cortical neurons to tones and noise in behaving macaque monkeys. Hear Res 150:104–118.

Recanzone GH, Guard DC, Phan ML (2000) Frequency and intensity response properties of single neurons in the auditory cortex of the behaving macaque monkey. J Neurophysiol 83:2315–2331.

Relkin EM, Pelli DG (1987) Probe tone thresholds in the auditory nerve measured by two-interval forced-choice procedures. J Acoust Soc Am 82:1679–1691.

Rolls ET, Treves A, Tovee MJ (1997) The representational capacity of the distributed encoding of information provided by populations of neurons in primate temporal visual cortex. Exp Brain Res 114:149–162.

Rosen MJ, Semple MN, Sanes DH (2010) Exploiting development to evaluate auditory encoding of amplitude modulation. J Neurosci 30:15509–15520.

Salinas E, Hernandez A, Zainos A, Romo R (2000) Periodicity and firing rate as candidate neural codes for the frequency of vibrotactile stimuli. J Neurosci 20:5503–5515.

Salinas E, Sejnowski TJ (2001) Correlated neuronal activity and the flow of neural information. Nat Rev Neurosci 2:539–550.

Schnupp JW, Hall TM, Kokelaar RF, Ahmed B (2006) Plasticity of temporal pattern codes for vocalization stimuli in primary auditory cortex. J Neurosci 26:4785–4795.

Schreiner CE (1998) Spatial distribution of responses to simple and complex sounds in the primary auditory cortex. Audiol Neurootol 3:104–122.

Scott BH, Malone BJ, Semple MN (2007) Effect of behavioral context on representation of a spatial cue in core auditory cortex of awake macaques. J Neurosci 27:6489–6499.

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

Shih JY, Atencio CA, Schreiner CE (2011) Improved stimulus representation by short interspike intervals in primary auditory cortex. J Neurophysiol 105:1908–1917.

Spitzer MW, Bala AD, Takahashi TT (2004) A neuronal correlate of the precedence effect is associated with spatial selectivity in the barn owl's auditory midbrain. J Neurophysiol 92:2051–2070.

Steinschneider M, Arezzo J, Vaughan HG Jr (1980) Phase-locked cortical responses to a human speech sound and low-frequency tones in the monkey. Brain Res 198:75–84.

Steinschneider M, Schroeder CE, Arezzo JC, Vaughan HG Jr (1994) Speech-evoked activity in primary auditory cortex: effects of voice onset time. Electroencephalogr Clin Neurophysiol 92:30–43.

Steinschneider M, Volkov IO, Fishman YI, Oya H, Arezzo JC, Howard MA 3rd (2005) Intracortical responses in human and monkey primary auditory cortex support a temporal processing mechanism for encoding of the voice onset time phonetic parameter. Cereb Cortex 15:170–186.

Suga N (1989) Principles of auditory information-processing derived from neuroethology. J Exp Biol 146:277–286.

Sutter ML, Margoliash D (1994) Global synchronous response to autogenous song in zebra finch HVc. J Neurophysiol 72:2105–2123.

Sutter ML, Shamma SA (2011) The relationship of auditory cortical activity to perception and behavior. In: The auditory cortex (Winer JA, Schreiner CE, eds), pp 617–643. New York: Springer.

Van Tasell DJ, Soli SD, Kirby VM, Widin GP (1987) Speech waveform envelope cues for consonant recognition. J Acoust Soc Am 82:1152–1161.

Victor JD (2000) How the brain uses time to represent and process visual information(1). Brain Res 886:33–46.

Walker KM, Ahmed B, Schnupp JW (2008) Linking cortical spike pattern codes to auditory perception. J Cogn Neurosci 20:135–152.

Walker KM, Bizley JK, King AJ, Schnupp JW (2011) Multiplexed and robust representations of sound features in auditory cortex. J Neurosci 31:14565–14576.

Werner-Reiss U, Porter KK, Underhill AM, Groh JM (2006) Long lasting attenuation by prior sounds in auditory cortex of awake primates. Exp Brain Res 168:272–276.

Woods TM, Lopez SE, Long JH, Rahman JE, Recanzone GH (2006) Effects of stimulus azimuth and intensity on the single-neuron activity in the auditory cortex of the alert macaque monkey. J Neurophysiol 96:3323–3337.

Yin P, Johnson JS, O'Connor KN, Sutter ML (2011) Coding of amplitude modulation in primary auditory cortex. J Neurophysiol 105:582–600.

Young ED, Barta PE (1986) Rate responses of auditory nerve fibers to tones in noise near masked threshold. J Acoust Soc Am 79:426–442.