Early Stages of Melody Processing: Stimulus-Sequence and Task-Dependent Neuronal Activity in Monkey Auditory Cortical Fields A1 and R

Pingbo Yin,1,2 Mortimer Mishkin,1 Mitchell Sutter,2 and Jonathan B. Fritz1

1Laboratory of Neuropsychology, National Institute of Mental Health, National Institutes of Health, Bethesda, Maryland; and 2Center for Acoustic and Auditory Research, Institute for Systems Research, University of Maryland, College Park, MD 20742 (E-mail: ripple@isr.umd.edu).

Submitted 16 September 2008; accepted in final form 4 October 2008

Yin PB, Mishkin M, Sutter M, Fritz JB. Early stages of melody processing: stimulus-sequence and task-dependent neuronal activity in monkey auditory cortical fields A1 and R. J Neurophysiol 100: 3009–3029, 2008. First published October 8, 2008; doi:10.1152/jn.00828.2007. To explore the effects of acoustic and behavioral context on neuronal responses in the core of auditory cortex (fields A1 and R), two monkeys were trained on a go/no-go discrimination task in which they learned to respond selectively to a four-note target (S+) melody and withhold response to a variety of other nontarget (S−) sounds. We analyzed evoked activity from 683 units in A1/R of the trained monkeys during task performance and from 125 units in A1/R of two naive monkeys. We characterized two broad classes of neural activity that were modulated by task performance. Class I consisted of tone-sequence–sensitive enhancement and suppression responses. Enhanced or suppressed responses to specific tonal components of the S+ melody were frequently observed in trained monkeys, but enhanced responses were rarely seen in naive monkeys. Both facilitatory and suppressive responses in the trained monkeys showed a temporal pattern different from that observed in naive monkeys. Class II consisted of nonacoustic activity, characterized by a task-related component that correlated with bar release, the behavioral response leading to reward. We observed a significantly higher percentage of both Class I and Class II neurons in field R than in A1. Class I responses may help encode a long-term representation of the behaviorally salient target melody. Class II activity may reflect a variety of nonacoustic influences, such as attention, reward expectancy, somatosensory inputs, and/or motor set and may help link auditory perception and behavioral response. Both types of neuronal activity are likely to contribute to the performance of the auditory task.

INTRODUCTION

A key question in auditory physiology is how complex sounds are encoded and how this sensory representation is linked to behavior. In the present study, we explored the neural representation of a target melody (four-note sequence) in core areas of auditory cortex of monkeys engaged in a simple auditory-discrimination task. Although higher-order representations of the complete target melody, or auditory object, may occur in auditory association cortex, we conjectured that such higher-order representations were likely to be based on simpler building blocks in the auditory core that encode component note frequencies and order (Kilgard and Merzenich 2002). Thus our experimental goals were to measure the response pattern to a highly learned target tone sequence to test whether neuronal responses to the target melody at the level of auditory core cortex in an awake, behaving animal were modulated by behavioral context and resolve whether the core neural representation of the melody could be parsimoniously explained as the linear sum of dyadic tone interactions between adjacent notes in the melodic sequence.

From earlier behavioral lesion studies, there is evidence that the auditory core plays an important role in the representation and discrimination of complex spectrotemporal patterns such as melody. Bilateral ablation of auditory cortex in cats impairs melody (tone pattern) discrimination (Diamond and Neff 1957; Dewson et al. 1970; Elliott and Trahiotis 1972; Neff et al. 1975; Tramo et al. 2002). Moreover, damage to the auditory cortex in humans is associated with impairments of discrimination of tone sequences (Pinheiro 1976), the order of tonal stimuli (Carmon and Nachshon 1971; Efron 1963; Swisher and Hirsh 1972), tone intervals and temporal rhythms (Liegeois-Chauvel et al. 1998), frequency change or direction discrimination (Tramo et al. 2002), and melodic contour (Peretz 1990). Imaging studies indicate that melody processing activates primary auditory cortex (A1) as well as secondary auditory association areas of the superior temporal gyrus and planum polare (Patterson et al. 2002).

In parallel, neurophysiological studies of the activity of the core or A1 in the behaving animal have been the focus of considerable current interest in the exploration of ongoing dynamic behavioral modulation of cortical responses and in the functional contribution of these changes to acoustic signal processing and behavior. In recent years, there has been considerable advancement in the understanding of neuronal response properties in A1 (reviewed in Eggermont and Ponton 2002; Ehret 1997; Nelken 2004; Read et al. 2002; Schreiner et al. 2000; Semple and Scott 2003). Earlier studies demonstrated that the activity and responses of some units in the core of auditory cortex depend on the attentional or behavioral state of the animal (Beaton et al. 1975; Benson and Hienz 1978; Benson et al. 1981; Fritz et al. 2003, 2005a,b, 2007a,b; Ho- churnan et al. 1976, 1981; Hubel et al. 1959; Miller et al. 1972; Pfingst et al. 1977; Vaadia et al. 1982) or by whether acoustic stimuli are externally presented or self-generated vocalizations (Eliades and Wang 2003, 2005, 2008). Previous work has also shown that A1 responses can be modulated by the history of previously presented sounds (Bartlett and Wang 2005; Brosch et al. 1997, 1998, 1999, 2000; Fishman et al. 2001, 2004; McKenna et al. 1998; Michely et al. 2005; O’Neill 1995; Ulanovsky et al. 2003, 2004; Wehr and Zador 2005; Werner-
Reiss et al. 2006; Zhang et al. 2005), suggesting that sensory context plays an influential role in shaping cortical responses, even in the anesthetized animal. Other studies have shown that A1 responses can be strongly modulated by auditory short-term memory (Gottlieb et al. 1989; Sakurai 1994) and influenced by long-term training on discrimination tasks (Bao et al. 2004; Beitel et al. 2003; Blake et al. 2002; Brosch et al. 2005; Recanzone et al. 1993). Behavioral training can lead to persistent memory traces and long-term changes in cortical organization (Kuroglinian and Knudsen 2007; Suga and Ma 2003; Weinberger 2004). Many, if not all, of these features of auditory core processing are likely to contribute to tone sequence and melody encoding.

We report results from recordings in two areas of the primary auditory cortex (fields A1 and R) of monkeys performing a go/no-go auditory-discrimination task. We analyzed the activity of neurons whose responses to a behaviorally salient tonal target (S+) sequence could not be predicted by their receptive field properties (as measured in response to single pure-tone transients). We reasoned that such task-dependent responses were likely to contribute to the cortical representation of salient auditory objects. We characterized two classes of neurons in the trained monkeys that showed nonlinear response patterns to the target (S+) melody.

Class I: The first class was comprised of sequence-selective neurons that showed short-latency highly facilitated responses or long-latency suppressive responses to tonal components of the four-tone target sequence. The short-latency facilitation was rarely observed in two naive monkeys in response to the target (S+) melody, suggesting that the short-latency facilitation observed in the two trained monkeys was indeed behavioral in origin. Similarly, the extent of suppression as well as the temporal pattern of suppressive responses in the trained monkeys differed from those observed in task-naive monkeys.

Class II: A second class of neurons was characterized by “anticipatory” activity that was highly correlated with reward and/or with task-related bar release, a behavioral response that could lead to reward.

These results suggest that the activity of auditory core neurons not only reflect past as well as present acoustic inputs but may also reflect a variety of nonacoustic influences, such as heteromodal sensory inputs, attention, reward, expectation, or motor set.

METHODS

Four adult male rhesus monkeys (Macaca mulatta), weighing 8–10 kg, were used. Experiments were performed in a double-walled sound-attenuating booth (IAC). All experimental procedures followed National Institutes of Health (NIH) guidelines. The protocol for experimental studies on the two trained monkeys (conducted at NIH) was approved by the National Institute of Mental Health animal care and use committee. The protocol for experimental procedures on the two task-naive monkeys (conducted at the University of California at Davis) was approved by the University of California, Davis animal care and use committee.

Behavioral task

All four monkeys were conditioned and habituated to being restrained in an acoustically transparent primate chair and head-fixed for physiological recordings. Two monkeys were trained to perform a go/no-go auditory-discrimination task.

Monkeys learned to grasp a touch bar mounted on the chair for 500 ms to initiate a trial. For target (S+) trials, the monkeys were required to release the bar within a 1.0-s behavioral response window to obtain a fluid reward (typically three drops of juice or water presented successively over a period of 600 ms). For nontarget (S−) trials, the monkeys were required to withhold release for ≥1.0 s after stimulus offset to avoid a 3.0-s timeout (Fig. 1B). Two slightly different behavioral response windows were used. For monkey B1, the 1.0-s response window began 250 ms after stimulus onset (S+ trial type 1 in Fig. 1B), whereas for monkey B2 it began from stimulus offset (S+ trial type 2 in Fig. 1B). For both monkeys, the likelihood ratio of the S+ stimulus presentations to S− stimulus presentations was 1:5 in random order.

Behavioral training

Behavioral training to the performance criterion (>80% correct for both S+ and S− trials for three behavioral sessions) required about 6

A Two variants of target (S+) melody

B Task paradigm

FIG. 1. Task schematic. The monkeys were trained to discriminate an S+ target melody (fixed duration of 500 ms) from a large set of acoustic stimuli, the nontarget (S−) sounds. The final nontarget (S−) stimulus set included single pure-tone transients, white and band-passed noise, frequency-modulated (FM) sweeps, and monkey vocalizations (with variable duration ranging from 50 to 800 ms). One of the monkeys (B1) learned only the low-frequency version of the S+ target melody and one monkey (B2) also learned a high-frequency version. A S+ target melody consisted of 4 sequential tones (1–2–3–1), illustrated as a color-coded bar (A). The frequencies of the component tones of the two S+ target melodies are given in A. Each of the first 3 tones had durations of 100 ms, whereas the duration of the last tone was 200 ms. A trial was initiated when the monkey held the touch bar for 0.5 s. A sound [either a target sound (S+) or nontarget sound (S−)] was then presented. In one task variant, one monkey (B1) was trained to release the touch bar for the S+ target melody within a 0.5-s response time window that began 250 ms after sound onset [B1: target sound (S+), trial type 1]. The second monkey (B2) learned another variant of the task, in which it was required to hold the touch bar throughout the target melody, and to release the touch bar immediately after offset of the target melody [B2: target sound (S+), trial type 2]. In nontarget sound (S−) trials, both monkeys were required to continue holding the touch bar for 1.0 s (B3). The monkey was rewarded for correct bar release on target sound (S+) trials with juice or water. Typically, the reward consisted of 3 drops, delivered one by one over a period of 600 ms. The intertrial interval for correct nontarget sound (S−) and target sound (S+) trials was about 1.5 s. The monkey received a 3.0-s timeout penalty for incorrect responses to target sound (S+) or nontarget sound (S−) trials (see text for details).
Acoustic stimuli: $S^+$ and $S^-$ sounds

$S^+$ sounds included only two variants, a high- and a low-frequency version of a four-note melody. Each of the two $S^+$ melodies was a sequence of four single pure-tone notes with the same frequency for the first and last notes. Tone duration was 100 ms for notes 1–3 and 200 ms for note 4, for a total $S^+$ duration of 500 ms (Fig. 1A). Although the onset and offset of the $S^+$ melody were ramped by 5-ms cosine rise/fall times (as were all $S^-$ stimuli as well), there was no gap or ramp between the individual four single pure-tone notes that comprised the melody. To reduce transients, each note used the final gap or ramp between the individual four single pure-tone notes that comprised the melody. To reduce transients, each note used the final gap or ramp between the individual four single pure-tone notes that comprised the melody. To reduce transients, each note used the final gap or ramp between the individual four single pure-tone notes that comprised the melody. To reduce transients, each note used the final gap or ramp between the individual four single pure-tone notes that comprised the melody. To reduce transients, each note used the final gap or ramp between the individual four single pure-tone notes that comprised the melody.

Physiological recording began when the $S^-$ stimulus set was comprised of 21 sounds, which included 10 rhesus monkey calls, 10 frequency-modulated (FM) sweeps (varied in sweep rates and directions), and white noise (500 ms). $S^-$ stimuli had durations from 50 to 800 ms. Eventually, in the final version of the task, the class of $S^-$ sounds included over 120 diverse stimuli divided into five sound sets: single pure tones, band-passed noise, AM or FM tones, and monkey calls. Since one of the key questions addressed in this study was whether auditory cortical responses to well-learned tone sequences (i.e., the $S^+$ melodies) could be predicted from the linear summation of responses to single pure-tone transients, the main set of $S^-$ sounds included 24 pure-tone transients, with evenly spaced (1/3 octave) frequencies from 100 to 20,300 Hz in pseudorandom order. This set included tones with matched frequencies near those of individual tones within the $S^+$ four-note melody. In this report, we describe neuronal responses to the pure-tone transient ($S^-$) stimuli and to the melody ($S^+$).

In addition to receiving the two $S^+$ sequences presented to the trained monkeys, the naive monkeys were also presented with other three-tone (four-note) control sequences (C-sequences) that were customized for each recording site based on the best frequency (BF) of the multineuron recording at that site (see Supplemental Material for further details).1 These additional, customized (target-like) stimuli were needed to elicit a sufficient number of facilitatory and suppressive physiological responses in the naive animals to allow comparison with the pattern of responses in the trained animals (see section entitled Comparison of enhancement and suppression patterns to four-note sequences ($S^+$ melodies and C-sequences) in trained and naive monkeys in Results). For the naive monkeys, all of the sounds were devoid of behavioral meaning or context, and random juice and water rewards were administered occasionally between passive sound presentation trials to keep the naive animals alert.

Presentation of acoustic stimuli

For all training and recording sessions, sounds were delivered free field via a single loudspeaker (trained monkeys: AP-7, Digital Phase; naive monkeys: Radio Shack PA-110) located in front of the monkey (trained monkeys: 0.75 m; naive monkeys: 1.5 m). Digitally synthesized sounds (16-bit output resolution) were calibrated and/or equalized, attenuated, and amplified so that the frequency output was relatively constant over the frequency range of 0.1–16.0 kHz (varying <5 dB). The two sound delivery setups were well matched in sound delivery (see Supplemental Material for further details). All experimental sessions were conducted in double-walled, sound-attenuating booths.

Surgical procedures and single-unit recording

After reaching the behavioral criterion, the trained monkeys were prepared for physiological studies. A titanium head-restraint post and a CILUX recording chamber (ICO-J2A, Crist Instruments) were implanted under aseptic conditions while the animals were deeply anesthetized with 2% isoflurane. Using the coordinates derived from MRI images, the recording chamber was centered on the dorsal (monkey B2) or lateral (monkey B1) surface of the skull, allowing access to the superior surface of the superior temporal gyrus.

Neurophysiological recordings were made using customized tungsten or glass-coated Pt/Ir microelectrodes (1.5–3.0 MΩ, FHC). Neuronal activity was filtered and amplified (TDT Bioamp system or A-M Systems Model 1800) and digitally acquired (POWER-1401, CED). Single-unit activity was isolated on-line and then confirmed off-line by spike-sorting software (SPSS or Spike2) that uses a template-matching algorithm. The primary auditory cortex was mapped in each monkey’s left hemisphere. Successive penetrations were separated by ≥0.5 or 1.0 mm in the mediolateral or rostrocaudal plane and recording sites in a given penetration were separated by ≥200 μm in depth. The electrode was advanced slowly and the depth of first auditory-responsive spikes was noted. We recorded from core auditory cortex at depths ranging from 100 to 2,000 μm below the first recorded auditory spike and thus we are likely to have recorded from most cortical layers (except possibly layer I and layer VI; see supplemental Fig. S1).

The physiological data from the two naive monkeys were gathered using similar recording techniques and with data-acquisition software developed at the University of California, Davis (O’Connor et al. 2005; Petkov et al. 2007).

Data analysis for Class 1 cells

Frequency-response areas (FRAs) were assessed using pure tones presented as $S^-$ stimuli. For trained monkeys, FRAs derived with $S^-$ tones during task performance were similar to those derived while not performing a task. The best frequency (BF) for the cell was taken to be either: 1) the frequency that elicited the largest response at the lowest intensity level tested or 2) the characteristic frequency (CF), if BF measurements were made at threshold intensity. To construct tonotopic maps, we used the median BF values of all units recorded from the same penetration to represent the BF for each cortical location.

Because monkeys B1 and B2 were highly trained to respond to the tones that comprised the $S^+$ melody, we used interpolation from responses to nearby $S^-$ tones to calculate a linear prediction of expected responses to those individual tones of the melody (Fig. 2). In detail, we derived the estimated poststimulus time histogram (e-PSTH) of the $S^+$ melody with the following procedure: We 1) calculated the PSTHs (with 10-ms bin size) for the two local frequency ranges, which included all of the frequencies used in the two $S^+$ melodies, and thus obtained a local FRA (Fig. 2A); 2) derived e-PSTHs for each of the individual three component tone-frequencies in each $S^+$ melody from the local FRA by using two-dimensional data interpolation (Fig. 2B); 3) estimated the PSTH (e-PSTH) for the complete $S^+$ melody by linear summation of the estimated responses to the individual components. From the expected responses to the pure-tone responses we predicted the time-varying response to the $S^+$ melody using linear summation of the expected tone responses. In this linear summation we included both estimated “on” and “off” responses (through 100 ms poststimulus) of the individual tonal com-

---

1 The online version of this article contains supplemental data.
ponents (Fig. 2C). This predicted PSTH (e-PSTH) for the S+ melody was then compared with the observed PSTH using a t-test (10-ms bin-by-bin basis).

An enhanced response to the S+ melody was defined as a minimum of two consecutive 10-ms bins in the observed PSTH with a significantly greater response ($P < 0.01$) than the predicted PSTH. A suppressed response was defined as at least two consecutive 10-ms bins in the PSTH with significantly less than the predicted PSTH. Thus the enhanced and suppressed responses were defined in relation to specific S+ notes rather than to the S+ melody as a whole. We also computed a strength index (SI) to quantify the magnitude of enhancement or suppression to the identified S+ melody notes. SI was defined as: SI = ($\bar{R} - \bar{Re}$)/($\bar{R} + \bar{Re}$), where R is the average response of two successive bins showing maximal activity in the observed response to the identified notes in the S+ melodies and Re is the average of the two corresponding bins in the estimated response. The SI varied from 0 to 1 and was applicable in cases of either suppression or enhancement. A low SI indicated minor differences between R and Re (i.e., SI = 0 when R = Re), whereas a high SI indicated major differences between R and Re and thus a highly suppressive or enhancement effect.

**Data analysis for Class II cells**

Each neuron was tested to ascertain whether its spiking activity was correlated with bar release. Initially, the neuronal activity associated with bar release was computed by temporally aligning all the correct S+ trials to the bar release, generating a bar-release–aligned histogram for a given behavioral block. Then bar-release timings were shuffled across trials and we recomputed the neuronal response by realigning the trials to the newly shuffled bar-release times. This shuffling procedure was repeated 100 times from which we obtained the average shuffled response and determined the 99% confidence interval. If the actual bar-release–aligned activity crossed the 99% confidence boundary of the shuffled neuronal activity within a 100-ms window before reward delivery (for monkey B1) or around bar release (for monkey B2), the activity was considered to be time-locked more strongly to the bar release than to the S+ sound. A neuron with such properties was classified as a Class II neuron. Two typical examples of Class II neurons, one from monkey B1 (Fig. 3A) and one from monkey B2 (Fig. 3B), are shown in Fig. 3. In the figure, the black curve indicates the actual neuronal responses aligned to bar-release times. The gray curve indicates the average shuffled neuronal response and the shaded area shows the 99% confidence interval.

We extended this method to extract the approximate nonacoustic component of activity from identified Class II neurons. As before, we first computed the average PSTH aligned to the S+ onset. We then subtracted this average acoustic PSTH from each individual trial of neuronal responses to remove the acoustic response evoked by the S+ sounds. Finally, we realigned each “acoustic-response-free” trial according to the timing of bar-release and thus obtained the activity evoked only by nonacoustic task components, such as motor command or hand tactile sensation of bar-release, attention, expectancy, or postbar-release reward delivery.

**RESULTS**

**Behavior**

Overall task performance for each trained monkey, averaged over all physiological recording sessions, is presented in Fig. 4A. Pooled performance level was high (>90% correct overall). However, both monkeys made incorrect go responses to a range of S− pure tones that were close in frequency to the individual notes comprising the S+ melodies (Fig. 4B). As shown in Fig. 4B, when the S− sounds were single pure tones, monkey B1 had an error rate of 49% for frequencies that most closely matched the first two notes in the S+(L) melody and 22% for the frequency that matched the third note of the S+(L) melody. For the same set of S− sounds, monkey B2 made fewer errors, responding incorrectly to an average of only 8 and 10%, respectively, of the frequencies that most closely matched the notes in the S+(L) and S+(H) target melodies.

This pattern of errors (particularly pronounced in monkey B1) suggests that one behavioral strategy that the monkeys used was a purely spectral “narrow-band, absolute-frequency” strategy to detect the presence of the S+ target sounds. This strategy relied on matching the pitches of individual notes to
those of the very similar component notes in the S+ melodies, rather than on matching to the template of the complete melody. This strategy was not surprising, since the initial S− sound set used in training consisted of nontonal sounds (either broader-band FM sweeps, white noise, or monkey calls). Thus for the initial training set, such a “narrow-band, absolute-frequency” strategy would be sufficient.

However, given the monkeys’ overall performance, this cannot have been the only strategy they used; otherwise, they would have shown an even higher error rate (i.e., >49% for B1 and >10% for B2) for S− pure-tone stimuli that closely matched the components of the S+ melody. Therefore both monkeys (particularly B2) presumably used an additional strategy to identify the S+ target melody, one that was sensitive to both its temporal and its spectral properties. The difference between the two monkeys in error rate may be due to the fact that monkey B2 was required to listen to the entire S+ melody before responding. This interpretation is supported by the

![(Image of graphs showing neuronal response PSTHs)](https://example.com/graph.png)

**FIG. 3.** Method for identifying Class II neurons. Neuronal response PSTHs were generated by aligning all correct trials to bar release (black curve) and to the shuffled bar-release time (gray curve; flanking gray-shaded area indicates the 99% confidence interval). The shuffling procedure is described in METHODS. In the 2 examples, neurons from monkey B1 (A) and monkey B2 (B) show a larger and sharper peak in actual PSTH than in the shuffled PSTH, thus showing that the neural activity was more closely time-locked to the bar release than to auditory stimulus onset. The vertical dotted lines indicate the bar-release time and the beginning of reward-delivery time (~200 ms in B).

![Image of graphs showing task performance](https://example.com/task_performance_graph.png)

**FIG. 4.** Task performance. A: average percentage correct performance is shown for each monkey (B1 and B2) during physiological recording sessions. B: S− (nontarget) error distribution shows the greatest likelihood of false-positive task errors at frequencies near those of the target melody S+ component tone frequencies. The peak nontarget S− error rate for matched frequencies reached nearly 50% for monkey B1, but was <10% for monkey B2 [when pooling performance from all behavioral blocks with both target S+ melodies, i.e., S+(L) and S+(H)]. C: distributions of reaction latency (bar-release likelihood) for each monkey for correct bar-releases to both S+(L) and S+(H) by monkey B2; the data for the 2 target melodies were combined because there was no observed difference between them in the distributions of bar-release latency. The vertical dashed line indicates the offset time of the S+ sound. D: distributions of reaction latency (bar-release likelihood) for each monkey for incorrect nontarget S− trials. Note similar timing to curves in C. The vertical dashed line indicates the offset time of the S− sound.
response-latency curves shown in Fig. 4C, in which the response-latency distribution was bimodal for monkey B1, with the major peak at 320 ms and a minor peak at 470 ms after sound onset. This latency distribution, with responses that began as early as 250 ms, was clearly shaped by the timing of reward, in that monkey B1 could obtain a reward by responding immediately after an obligatory 250-ms waiting period following the onset of the S+ sound. The response latencies for monkey B2 were longer, with a peak at 600–620 ms, reflecting the fact that this monkey was required to wait until the S+ melody finished (500 ms) before responding. For comparison, the distributions of response latency for erroneous releases to single S− tones are shown in Fig. 4D. These error latencies are similar to the correct response latencies for both monkeys, as shown for the correct responses to the S+ in Fig. 4C.

Overview of the physiological recordings

This neurophysiological report is based on a comparison of the extracellular responses of 683 single neurons from multiple depths in primary auditory cortex (A1 and R) of two trained monkeys performing the auditory-discrimination task with the extracellular responses of 125 single neurons in primary auditory cortex to similar acoustic stimuli in two naive monkeys. The comparison is limited to neuronal responses to the S+ and S− sounds. (Responses to the other S− sound sets—FM sweeps, band-passed noise bursts, and monkey calls—will be the subject of future reports.)

Overall, we found that neuronal firing was strongly modulated during presentation of the S+ melodies in the majority of units in trained monkeys, with 22% (135/615) showing enhancement and 67% (412/615) showing suppression (Fig. 5, white bars). Such widespread modulation to the same four-note sequence was not observed in recorded units of naive monkeys, which showed only 3% (2/73) enhancement and 67% (412/615) showing suppression (Fig. 5, gray bars). Thus there was an eightfold increase in enhancement effects to the S+ melody in the trained monkeys compared with that found in the naive monkeys and a nearly twofold increase in suppressive effects to the S+ melody in the trained compared with the naive monkeys.

In detail, the neurons in the trained monkeys were recorded in the course of 37 electrode penetrations (23 in N1, 14 in N2). The distribution of neuronal BF tuning in the naive monkeys overlapped that in the trained monkeys across five octaves (Supplemental Fig. S2), yet, as indicated earlier, they showed far less modulation during presentation of the S+ melodies.

In addition to the S+ melodies, a set of customized four-note sequences centered at neuronal BF was also presented to the task-naive monkeys (see METHODS). As mentioned earlier, the rationale for using these additional, customized (target-like) stimuli was to elicit a sufficient number of enhanced and suppressed neuronal responses in the naive animals to allow comparison with the pattern of neuronal responses in the trained animals. Hence, in the naive monkeys, all 125 units were studied with a set of six 1/3-octave customized-sequence stimuli (C-sequences). As expected, we observed a greater proportion of enhancement and suppression in these neurons than that in those studied earlier with a single set. Overall, 23% (29/125) showed enhancement and 58% (72/125) showed suppression to at least one of the six permuted C-sequences that were constructed by using the frequencies around each neuron’s BF (see METHODS).

Frequency maps

Recording sites covered much of the core auditory cortex and included a frequency reversal at the A1 and R border. This is shown in the grayscale plots (Fig. 6, bottom panels) of the BF maps in the two trained monkeys, ranging from low frequency (light gray) to high frequency (black). In both trained monkeys, the tonotopic axis in A1 ran from caudomedial to rostrolateral, with the highest BFs located caudally and the lowest BFs located rostrally. Continuing rostrally, there was a reversal of the tonotopic organization in both monkeys, consistent with the mirrored tonotopic organization in area R relative to that in A1.

The BFs in fields A1 and R in the trained monkeys showed less variation in the mediolateral direction than that in the rostrocaudal direction. This tonotopic organization is consistent with the isofrequency organization of A1 and R described in anesthetized macaque monkeys (Hackett et al. 2001; Kaas and Hackett 2000; Kosaki et al. 1997; Merzenich and Brugge 1973; Rauschecker et al. 1997) and in awake macaque monkeys (Recanzone 2000). In monkey B1, there was also a second reversal of the tonotopic axis rostral to R, which may mark the
TABLE 1. Summary of distribution of Class I neurons in trained and naive monkeys

<table>
<thead>
<tr>
<th></th>
<th>(+)</th>
<th>(−)</th>
<th>(+) and (−)</th>
<th>No Effect</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trained (n = 2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S+(L)</td>
<td>92</td>
<td>371</td>
<td>64</td>
<td>216</td>
<td>615</td>
</tr>
<tr>
<td>S+(H)</td>
<td>63</td>
<td>186</td>
<td>44</td>
<td>163</td>
<td>368</td>
</tr>
<tr>
<td>S+(L) or S+(H)</td>
<td>135</td>
<td>412</td>
<td>98</td>
<td>166</td>
<td>615</td>
</tr>
<tr>
<td>Naive (n = 2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S+(L)</td>
<td>1</td>
<td>15</td>
<td>1</td>
<td>58</td>
<td>73</td>
</tr>
<tr>
<td>S+(H)</td>
<td>1</td>
<td>13</td>
<td>1</td>
<td>60</td>
<td>73</td>
</tr>
<tr>
<td>S+(L) or S+(H)</td>
<td>2</td>
<td>26</td>
<td>2</td>
<td>47</td>
<td>73</td>
</tr>
</tbody>
</table>

(+) enhancement; (−), suppression. See text for further explanation.

boundary with the rostrotemporal field (area RT), one of the three auditory core areas defined in the primate (Bendor and Wang 2008; Kaas and Hackett 2000; Morel and Kaas 1992).

Although less densely sampled, recording sites in the two naive monkeys were also centered on the core auditory cortex (Supplemental Fig. S3). Based on the tonotopic maps, the recordings in monkey N1 were from A1 and those in N2 were from A1 and R.

Enhanced responses to S+ melodies in trained monkeys

As noted earlier, the responses to the frequencies in the S+ melodies were not always predicted by the linear estimation of the responses to the component single tones alone (e-PSTH). An enhanced response was found to the frequencies in the melodies S+(L) or S+(H) in 22% (135/615) of the neurons (Table 1, column 2). Enhancement could be striking. Even when there were virtually no responses observed to the single tones presented alone, there could be a large response evoked by the melodies in the same frequency region as the component tones. The enhanced responses evoked by the S+ sound occurred at any one or more of the four note positions in the melodic sequence. Individual examples of enhancement to tone components in the melody, from six different neurons, are given in Fig. 7. A population profile of enhancement is given in Fig. 10.

As shown in Fig. 7, the enhanced responses evoked by the S+ melody could occur at either the second (Fig. 7A, C, E, and F), third (Fig. 7B–F), or fourth note (Fig. 7E and F) in the melodic sequence. In three neurons, there were enhanced responses to only one note in the melody (Fig. 7A, B, and D), whereas the other three neurons showed enhanced responses to multiple notes in the melody (Fig. 7C, E, and F). Two neurons showed enhanced responses to all three notes (Fig. 7E and F). In more detail, the unit in Fig. 7A showed no response at all to any tones in the frequency range from 317 to 1,270 Hz presented as the S− single tones, but responded vigorously with an enhanced response to the second note of the S+(L) target melody, which had a frequency of 665 Hz (in the middle of the S− pure-tone unresponsive frequency region). The neuron in Fig. 7B showed only a weak off-response to S− pure tones in the frequency range of 800–1,008 Hz, although the note in S+(L) melody with a frequency of 790 Hz evoked an enhanced on-response and a strong, sustained response that was maintained for the full duration of the note. The neuron in Fig. 7C displayed a sustained response to S− pure tones in the frequency range <400 Hz and did not respond at all to tones with frequencies >400 Hz when presented alone as S− stimuli. This neuron, however, responded strongly with marked enhancement to the second and third notes of the S+(L) melody (i.e., 665 and 790 Hz). The neuron shown in Fig. 7D responded to S− tones in the frequency range of 4,032–12,800 Hz with onset or weak sustained responses, with the exception of the S− tone at 8,063 Hz, to which the cell did not respond at all when it was presented as an S− tone. If anything, there may have been a suppressed response to this S− tone. However, as shown in Fig. 7D, there was a strongly enhanced response to the third note of the high-frequency target melody S+(H), which had a frequency of 8 kHz, very close to the frequency that had elicited virtually no response when presented as an S− sound. In Fig. 7E, there was a very strong
onset response evoked by the second, third, and fourth notes of the S+(H) melody, even though there was little or no response to any S− tone except for a very small onset response at 8,063 Hz. This neuron also offers an interesting example of differential responses to the same frequency tone (5,000 Hz) located in different positions of the target melody (first and last positions). Although there was no response or faint response to the first note, strongly enhanced activity occurred in response to the fourth note, a sequence position that made it more clearly part of the S+ melody. The final example is shown in Fig. 7F. The responsive frequency range for this neuron to S− tones was narrow, restricted to an onset response to a single frequency (5,080 Hz). However, the neuron showed a strongly enhanced response to the third note (8 kHz) of the S+(H) target melody. As in the earlier case (Fig. 7E), this neuron too showed an enhanced response to the fourth note in the S+(H) melody (5 kHz), although the response to the first note was predictable from the S− tone response (Fig. 7, E and F). These examples illustrate striking enhancement to the notes in the S+ melody and also the selective nature of the enhancement—i.e., not only to the frequency of a given component note, but also to its position in the melodic sequence.

**Suppressed responses to target (S+) melodies**

We also observed sequence-selective suppressive effects in many neurons, e.g., neurons that responded well to single pure-tone transients when presented as S− stimuli but showed a relatively diminished response when the same frequencies were presented as notes in the four-tone S+ melodies. As shown in Table 1, in the trained animals there was a threefold greater likelihood of suppressive effects [67% (412/615) to S+(L) or S+(H) melodies] than enhancement effects [22% (135/615)]. Individual examples of sequence-specific suppression from three neurons are illustrated in Fig. 8 and a population profile of suppression is shown later in Fig. 10.

The two neurons shown in Fig. 8, A and B did not respond at all to the S+(L) melody, even though they showed a marked response to the S− tone with a frequency of 635 Hz, close to the frequency of the second note (665 Hz) of the S+(L) melody. Another example of suppression is shown in Fig. 8, C and D. This neuron had at least two responsive frequency areas (317–1,008 and 4,032–12,800 Hz, Fig. 8, C and D, respectively) that correspond to the frequency range of the component notes in the low- and high-frequency S+ melodies. Although this unit showed a sustained/prolonged response to a wide range of S− tones, it showed profound suppression to the S+(L) melody (i.e., no response at all) and also showed a parallel, suppressed response to the S+(H) melody.

**Comparison of enhancement and suppression patterns to four-note sequences (S+ melodies and C-sequences) in trained and naive monkeys**

In recording from primary auditory cortices in the two naive monkeys, we found only 2 of 73 units that showed response enhancement to either of the S+ melodies. Using C-sequences, however (see METHODS), the enhancement percentage increased eightfold to about 23% (29/125) from <3% (2/73) and was thus comparable to the enhancement distribution seen in the trained animals. The construction of the C-sequences optimized the likelihood of obtaining context-dependent response modulation in the naive animals, since we 1) customized the C-sequences to be close to each neuron’s BF and 2) included modulations to all six possible combinations of notes (rather than simply to the 1–2–3–1 sequence). Had we correspondingly modified the task and customized the S+ melody for each recording in the trained animal, or included all modulations for all note combinations, it is likely that we would have observed a much higher prevalence of modulation effects in the trained animals.

However, the modulation effects obtained from the C-sequences in the naive animals were useful for comparison with modulation observed to S+ melodies in the trained monkeys. We first analyzed the data to determine whether there were any
FIG. 8. Class I neuronal target melody $S^+$ suppression profiles. Suppressed responses of 3 neurons to individual tone components in the low- and high-frequency $S^+$ target melodies. The panel layout is the same as described in Fig. 7. The blue downward arrows indicate the bins showing suppression. A and B: 2 neurons showed suppression to the $S^+(L)$ target melody, particularly to the 2nd note. Although both narrowly tuned neurons displayed a strong response to the single nontarget $S^-$ tone frequency of 635 Hz (bottom rasters in A and B), neither responded to the component notes of the $S^+(L)$ target melody at nearby frequencies such as the 2nd tone frequency of 665 Hz. In addition to strong suppression of the 2nd tone frequency in the $S^+(L)$ target melody, the 2nd neuron (B) also showed significant (though less) suppression to the 3rd and 4th tone frequencies in the $S^+(L)$ target melody. C and D: responses of a 3rd neuron both to the $S^+(L)$ target melody (C) and also to the $S^+(H)$ melody (D). In this broadly tuned neuron, individual nontarget $S^-$ single pure tones evoked strong and sustained responses to tones over a wide frequency range that encompassed all component frequencies in both target melodies $S^+(L)$ and $S^+(H)$. However, as shown in C, there were virtually no responses to the target melody $S^+(L)$, and thus marked suppression. Also, as shown in D, there was a strongly suppressed response to the 2nd tone of the $S^+(H)$ target melody, although there was a clear response to the 3rd and $S^+(L)$ component note in the high-frequency target melody.

differences between the two populations in the strength of enhancement or suppression in response to the tone sequences, using the strength index (SI; see Methods). The median SI value for the enhancement effect to the $S^+$ melodies in the neuronal population from the trained monkeys (0.723, $n = 190$) was greater than that observed in naive monkeys (0.43, $n = 2$) and it was also greater than the median SI value of 0.603 ($n = 72$; $P < 0.01$, Wilcoxon rank-sum test) for the 1/3-octave C-sequences presented to the naive monkeys. The comparison is shown in Fig. 9A1 (top).

A similar result was obtained when we compared the magnitude of the suppression effect in the two neuronal populations. The neurons that exhibited suppression in the trained monkeys showed a significantly larger median SI value (0.823, $n = 1,107$) than the SI value found in the task-naive monkeys both for $S^+$ melodies (0.775, $n = 28$, $P < 0.001$) and for C-sequences (0.685, $n = 316$, $P < 0.005$, Wilcoxon rank-sum test for both). This comparison is shown in Fig. 9B1 (top).

We also found that the temporal pattern of the enhanced responses in the naive and trained monkeys was strikingly different. In the task-naive monkeys, most of the enhancement
responses occurred during the latter half of a given note in the C-sequence (i.e., with a latency >50 ms). This observation is compatible with the results from previous research studying responses to tone pairs in naive monkeys (Bartlett and Wang 2005; Brosch et al. 1999), which found that in most neurons, the enhanced responses to the second tone occurred during a late response phase after the initial response. In sharp contrast, the enhanced response in our trained monkeys had a much shorter latency, occurring during the onset response to a given note in the S+ melody.

To quantify the temporal differences between enhanced (and suppressed) responses in the trained and naive animals, we directly compared the time course of the neuronal population responses from the two pairs of animals. We computed probability density functions (PDFs) on a 10-ms bin-by-bin basis over a time course spanning the second, third, and fourth notes of the tonal sequences. For this analysis we used 400 bins from 135 neurons showing enhanced responses in the trained monkeys and 221 bins from 29 neurons with enhanced responses in the naive monkeys. As shown in Fig. 9A2 (middle), the likelihood of a bin showing an enhanced response in the trained monkeys peaked at 25 ms (corresponding to the third bin 20–30 ms following note onset within the S+ melody). In contrast, the likelihood of an enhanced response in the naive monkeys peaked much later, at 75 ms (corresponding to the eighth bin 70–80 ms following note onset within the C-sequences). The two-sample Kolmogorov–Smirnov (KS) test showed a highly significant difference (P < 0.001) in the distributions of the temporal occurrence of enhancement in the two neuronal populations. Figure 9A2 shows this difference in time course of enhancement between trained and naive monkeys.

Figure 9B2 shows the same analysis applied to suppressive effects. In the trained monkeys, the population density functions were constructed from bins with suppressive effects to the S+ melodies (a total of 5,373 bins from a total of 412 neurons; gray curve in the figure). In the naive monkeys, the population density functions were constructed from bins showing suppression to the C-sequences (a total of 1,433 bins from 72 suppression neurons; black curve). As with the enhancement effects, there were significantly different patterns in the time course of the suppressive response in the trained and naive monkeys. The PDF from the naive monkeys showed an early sharp peak (at 25 ms, corresponding to the third bin, 20–30 ms from note onset), whereas the peak of the PDF for the trained monkeys was relatively broader and occurred significantly later in time (at 35 ms, corresponding to the fourth bin, 30–40 ms from onset). As earlier, the two-sample KS test showed a highly significant difference (P < 0.001) between the distributions of the temporal occurrence of suppression in naive and trained monkeys. Figure 9B3 shows this difference in time course of suppression between the two groups.

The “heat” maps in Fig. 10 show the population profiles of enhancement effects (in red) and suppression effects (in blue) across neurons and time course. The heat maps are sorted in the order of the timing of the first modulated (enhanced or suppressed) 10-ms bin for individual neurons. The maps reveal several interesting differences in enhancement and suppression response patterns between trained and naive monkeys. First, the onset times of the modulated bins (for both enhancement and suppression) are more sparsely distributed in neuronal responses of the trained monkeys than in those of naive monkeys. Also, in accord with the results shown in Fig. 9, in naive monkeys, enhancement effects (to C-sequences) typi-
cally occurred with a relatively long latency after each of the note transitions or occurred following a prior wave of suppressive effects (Fig. 10B). The amplitude of these enhancement effects in the naive monkeys was comparatively weak. Thus the overall pattern, shown in Fig. 10B (rightmost panel), reveals an alternating pattern of successive waves of suppression and enhancement for each note transition in the four-note C-sequence. At a global level, this pattern did not vary much between successive note transitions. In marked contrast, a very different overall population pattern occurred in the responses of trained monkeys, since the enhancement effects showed greater amplitude and occurred with a shorter latency, at about the same time as the suppressive effects (Fig. 9). The net effect of this differential timing in responses in the trained animal was to form a striking “coincidence wave” of modulation that continuously progressed across the neuronal population from target (S+) melody onset to offset and was followed by periods of subsequent suppression (Fig. 10A). This moving “coincidence wave,” beginning with the response to the first note, may encode the occurrence of the initial, salient S+ “event” and the subsequent tone transitions that characterize the S+ melody.

**Nonauditory neuronal activity: Class II**

In analyzing neuronal activity during task performance, we observed neurons that showed nonauditory activity that was highly correlated with 1) task-related bar release, the behavioral response that could lead to reward, and/or 2) task-related onset of bar grasping, and/or 3) the anticipation of reward. We also observed neurons that responded during reward delivery; however, since there were unavoidable, multiple acoustic features associated with reward delivery (clicking of solenoid for water release, licking of the monkey) we focused our analysis instead on reward-related activity that preceded bar release and actual reward delivery (see METHODS for description of criteria for classification as Class II neurons). Out of a total of 683 single neurons isolated during the performance of the task, 13% (90/683) showed pre-reward-related activity (see Table 2). Of these 90 units, 20 cells from monkey B1 and 65 from monkey B2 showed bar-release–related activity. Moreover, 13/70 cells from monkey B2 showed a separate and distinct increase in activity prior to reward delivery (which closely coincided in time with the onset of bar grasping).

The neuronal responses shown in Fig. 11 (A and B) were recorded from two neurons in monkey B1. This monkey had been trained to release the touch bar as soon as possible after detecting the S+ melody for immediate reward following a correct bar release (Fig. 1B1, S+ trial type 1). In this behavior, the quickest bar release took ≥ 250 ms from the onset of the S+ and bar-release probability peaked at 320 ms following S+ onset. Neuron 1 (Fig. 11A) had a predictable transient on response to the S+ and a striking long-lasting response that began at about 300 ms (Fig. 11A, left). This activity was only loosely correlated with any feature of the S+ melody. However, by realigning all the spikes in relation to bar release and delivery of the reward (Fig. 11A, right), a much more tightly correlated pattern emerged. Spike activity began at about 50 ms before the bar release or the delivery of reward (see blue arrow). A similar pattern was observed in the spiking activity of neuron 2 (Fig. 11B). Although neuron 2 was tuned to a high frequency at around 12.8 kHz, it had a response evoked by the S+(L) melody that began about 300 ms after S+(L) onset (Fig. 11B, left). As in neuron 1, the activity in neuron 2 also showed a better correlation to the upcoming bar release or reward delivery (Fig. 11B, right). An increase in activity in neuron 2 appeared about 50 ms before bar release or reward delivery. To ascertain whether this activity was related to bar release or to anticipation of reward, we delayed the delivery of the reward to 100 ms after a correct bar release (as shown in Supplemental Fig. S4). Spikes from neuron 2 showed a better correlation with bar release than with reward delivery, thus suggesting that this spiking activity component might be related either to tactile sensation or motor output during bar release. In contrast, the neuron shown in Fig. 11C from monkey B2, to which reward was delivered 200 ms after the bar release rather than simultaneously with it (Fig. 1B2, S+ trial type 2), shows clear postbar-release–related activity as well as activity in anticipation of reward (Fig. 11C, right). However, in general, bar-release–related activity was not as strong or as long-lasting in monkey B2 as in monkey B1 (Fig. 11A), even though the activity of 65 neurons in monkey B2 was of this type. This physiological difference may be related to the distinct task variants learned by each monkey [see METHODS (Fig. 1B) and also DISCUSSION].

The firing patterns of the three individual Class II neurons shown in Fig. 11 clearly illustrate activity related to nonacoustic events. To provide an overview of the time course of all Class II activity in relation to these events, we summarize population activity in Figs. 12 and 13 (from monkeys B1 and B2, respectively).

The population activity of Class II neurons from monkey B1 was aligned to stimulus onset and showed an initial transient onset response to the S+ melody (Fig. 12A). All Class II neurons recorded in A1 or R showed auditory responses of this kind. Following the sensory onset response, there was a gradual increase leading to a second, strong, long-lasting period of activity that occurred about 250 ms after stimulus onset (Fig. 12A). This second, long-lasting period of activity, observed in correct trials, was strikingly shortened and diminished in the population spiking activity from incorrect bar-release trials to S− tones (Fig. 12C). After realigning the population activity for all the experimental trials to bar release/reward delivery, we observed a large peak (blue arrow) during correct trials that had a clear buildup phase (Fig. 12B, red area). In contrast, during incorrect bar release to S− tones, relatively weak bar-release activity was observed (Fig. 12D). This population pattern of auditory cortical activity, preceding and following bar release, illustrates the contribution of nonacoustic inputs in modulating

**TABLE 2. Class II neurons showing nonauditory responses during task performance**

<table>
<thead>
<tr>
<th>Monkey</th>
<th>Total Class II Neurons</th>
<th>Bar Release–Related Response</th>
<th>Reward–Related Response</th>
<th>Total Neurons Analyzed</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1</td>
<td>20 (15.6%)</td>
<td>NA*</td>
<td>NA*</td>
<td>128**</td>
</tr>
<tr>
<td>B2</td>
<td>70 (12.6%)</td>
<td>65</td>
<td>13 (8 overlap)</td>
<td>555**</td>
</tr>
</tbody>
</table>

*Since the rewards were delivered immediately after the bar release in monkey B1, it was unclear whether the response was associated with bar release or with anticipation of reward. **This number includes 68 neurons that were excluded from Table 1 because they had been tested not with the S− tone set but with the other S− sound sets instead.

**J Neurophysiol • VOL 100 • DECEMBER 2008 • www.jn.org**
Class II neuronal activity. This population activity is in accord with results from single neurons shown in Fig. 11.

When population activity for correct trials of monkey B2 was aligned to stimulus onset, it showed a small acoustic onset response to each of the four notes of the S+ melodies (initial four gray peaks in Fig. 13A) as well as a long-lasting buildup response following offset of the melodies. A large transient peak following bar release (purple arrow) became visible when the activity was realigned to bar release (as shown in Fig. 13, B and D). This peak of activity associated with bar release was preceded by relatively weak activity (Fig. 13, B and D). As in monkey B1, the population response during correct bar-release trials was greater than the response from incorrect bar release to S− tones (Fig. 13, C and D).

Additional evidence for reward-related activity is given in Fig. 14, which shows the population activity from a small subset of Class II neurons (n = 13 from monkey B2). The histograms in this figure, aligned to stimulus onset (Fig. 14A) and to bar release and reward delivery (Fig. 14B), illustrate a pattern of gradual buildup activity prior to reward delivery (green arrow) that was clearly distinct from the bar-release response (purple arrow). As mentioned, one important caveat is that bar regrasp occurred close to the time of reward delivery, so that it is difficult to disentangle the contributions to this activity of motor, cognitive, and nonauditory sensory inputs.

By extracting the nonacoustic responses of Class II neurons (see METHODS) and pooling these responses for each monkey, it was possible to obtain population profiles of Class II effects. As shown in the two panels of Fig. 15, the population responses from identified Class II neurons in the two trained monkeys had similar response patterns: suppression began at approximately 250 ms and reached a valley at 50–100 ms before bar release, which was followed by an excitatory peak near the bar release.

In conjunction with observational evidence that neither of the trained monkeys was vocalizing, nor making task-realted sounds during task performance (see Supplemental Fig. S5), these individual neuronal and population responses of bar-release–related activity provide strong evidence for nonacoustic activity in the auditory cortical core of the macaque monkey.

**Differential distribution of Class I and Class II neurons in A1 and R**

We analyzed the location of Class I and Class II neurons in the auditory core areas (Fig. 16), based on the mediolateral position of the penetrations relative to the tonotopic maps (Fig. 6). In field A1, 66% (292/441) of the cells were classified as Class I and 11% (47/441) were classified as Class II. By
DISCUSSION

Until recently, the function of the primary auditory cortex has been viewed mainly from the perspective of auditory information processing. However, new evidence suggests that activity in A1 and R can also be modulated by top-down influences such as attention and expectation (Fritz et al. 2007a,b) and by other nonacoustic influences (Forte and Schroeder 2005; Ghazanfar and Schroeder 2006). In these experiments, we observed the presence of neuronal activity that was dependent on behavioral as well as acoustic context and, accordingly, we have defined two new classes of neuronal properties: 1) Class I: sequence selective and 2) Class II: nonauditory. Since we found neurons with these properties at all cortical depths, our data do not provide evidence for laminar segregation of these two classes (see Supplemental Fig. S1). However, our data do indicate significant quantitative differences in the proportion of these classes in A1 and R (Fig. 16).

We speculate that the Class I responses may comprise a set of mechanisms that constitute an early stage of encoding of a behaviorally salient melody and may contribute to a long-term sensory representation of the melody. These mechanisms offer insight into the neural correlates in the auditory cortical core of melody processing for a very simple tone sequence. In contrast, we propose that Class II activity may help link auditory perception and behavior (since all type II neurons also responded to acoustic stimuli). Both types of activity are likely to contribute to performance of audiomotor tasks and illustrate the role of the auditory core in developing an early cortical representation of salient acoustic stimuli and a unified representation of auditory as well as relevant nonacoustic events.

In the following two sections we examine our results showing Class I enhancement and suppressive effects to tone sequences in the context of previous studies. In the next two sections, we discuss these sequence-related response modulations in terms of possible behavioral strategies used by the trained monkeys and the implications for representation of auditory sequences in A1 and R. Finally, we discuss our observations of Class II nonauditory responses in the context of other recent findings of multisensory representations in primary sensory cortices.

FIG. 12. Population analysis of Class II nonacoustic responses; evidence for prebar-release activity. The population PSTHs (gray histograms in A–D) represent the average of individual normalized PSTHs (one from each neuron in the population). The PSTH from each individual neuron was normalized to the mean spontaneous baseline, which is indicated by a horizontal dot-dashed line in each population plot. The black histograms in A and C show the time distribution of bar release during task performance and the vertical dot-dashed lines mark the peak time of the distributions. A and B: population response (20 neurons) of correct bar-releases to S+ melodies, aligned to the onset of these stimuli (A) and to bar-release (B), respectively. A: this population PSTH was derived from responses of Class II neurons in monkey B1 during performance of the type 1 version of the task, which permitted bar-release immediately after onset of the S+ melody followed by immediate reward delivery. Note responses to onset of the S+ stimulus as well as to bar-release and reward delivery (seen more clearly in B, where the same spike data are aligned to bar-release). The red portion of the release-aligned histogram (B) indicates premovement activity (50 ms prior to bar-release) and water delivery time (indicated by the green dotted line). C and D: population activity of incorrect bar-release to the nontarget S− stimuli. The population PSTH in C was aligned to onset of the acoustic stimuli and in D it was aligned to bar-release. Unlike the case in B for correct release to the S+ melody, in D there is relatively little buildup to bar-release for incorrect releases to S− sounds.

contrast, in field R, 86% (150/174) of the cells were classified as Class I and 23% (40/174) as Class II (16 neurons had both Class I and Class II properties). These interareal distributions were significantly different using the chi-square test ($P < 0.001$ for both Class I and Class II).
Class I effects and tone-sequence enhancement of responses in auditory cortex

Some of the earliest studies to examine facilitated responses to pairs of stimuli (Feng et al. 1978; O’Neill and Suga 1979) observed neurons in the auditory system of echolocating bats that responded selectively with enhanced responses to pairs of tones or FM chirps at specific interstimulus delays between acoustic stimuli, corresponding to behaviorally relevant echo delays. In nonecholocating species such as cat and monkey, enhanced responses to two-tone combinations have been demonstrated (Bartlett and Wang 2005; Brosch and Schreiner 1997, 2000; Brosch et al. 1999). A high proportion of neurons in primate primary auditory cortex exhibited intrinsic facilitation to two-tone combinations if the appropriate tone pairs were presented (Bartlett and Wang 2005; Brosch and Schreiner 1997, 2000; Brosch et al. 1999; De Ribaupierre et al. 1972). To elicit maximal facilitation the tone pairs had to be chosen carefully and often required highly specific spectral and temporal separation of the tone-pair stimuli (Bartlett and Wang 2005; Brosch and Schreiner 2000). Brosch and colleagues (1999, 2000), in recordings from auditory cortex in anesthetized macaque monkey and cat, found that the combination of two tones (T1 and T2) most likely to elicit greatest T2 facilitation is composed of tone pairs separated by about 1.0 octave in frequency and about 100 ms in time of onset. There are even fewer studies of sequences composed of three or more tones (Espinosa and Gerstein 1988; McKenna et al. 1989). Earlier studies as well as the present study generally find context dependence that is not predictable from responses taken outside of the context of the sequence.

![Diagram](http://jn.physiology.org/)

**FIG. 13.** Population analysis of Class II nonauditory responses; evidence for postbar-release activity. Population PSTH and release-aligned histograms (in gray) from Class II neurons (n = 65) in monkey B2 during performance of the type 2 version of the task, which required delaying bar-release until offset of the S+ melody, with reward delivery occurring 200 ms later. A and B: population activity for all correct behavioral responses (bar-release after S+ melodies) aligned to stimulus onset (A) and to bar-release (B). C and D: population response for all incorrect bar-releases to S+ stimuli. The red portion (width = 50 ms) of the bar-release-aligned histogram (B) indicates an increase in activity that begins ≥50 ms prior to bar-release (as can be seen in B the increase in activity begins ~150 ms before bar-release). However, as clearly shown in B and D, there is much greater activity following bar-release than before bar release [not only for correct responses (B) but also for incorrect responses (D)]. The green dotted line in B marks the water reward delivery onset time and the green portion (duration = 50 ms) of the PSTH (B) indicates an absence of anticipatory activity prior to water reward. The PSTH in B indicates that most of the increased activity occurred after the onset of water delivery. The black histograms in A and C show the temporal distribution of bar-release during task performance.

**FIG. 14.** Population analysis of Class II nonauditory responses; evidence for pre-reward activity. Population PSTH and release-aligned histograms (in gray) from Class II neurons (n = 13) with responses highly correlated with reward delivery. All responses were from monkey B2 during performance of the type 2 version of the task. The green dotted line in B indicates the water delivery time (200 ms after a bar-release). The red portion (duration = 50 ms) of the release-aligned histogram (B) indicates pre-movement activity that begins ≥50 ms prior to bar-release. The green portion (duration = 50 ms) of the release-aligned histogram (B) indicates anticipatory activity that begins ≥50 ms prior to water reward. The black histogram in A shows the temporal distribution of bar-release during task performance. Note the presence of an auditory response to the S+ melody in A and, in B, a small response preceding and following bar-release, as well as strong activity time-locked to the start of reward delivery.
Although there are a handful of neurophysiological studies that have investigated cortical responses to two-tone combinations, there are even fewer that have explored responses to longer tone sequences. Although less frequently studied, sequence-sensitive neurons have been described in the auditory cortex of the cat and monkey (Espinosa and Gerstein 1988; McKenna et al. 1989; Yin et al. 2001). The study by McKenna and colleagues (1989) explored responses of neurons in cat primary and secondary auditory cortex to a variety of tone sequences (each composed of five isointensity 300-ms tones with 300- to 900-ms intertone intervals) and demonstrated that the response to the tones within the sequence was not predictable from the responses to isolated tones (presented with 3,200-ms intertone intervals). Espinosa and Gerstein (1988) reported order sensitivity in neural networks in A1 by showing that cross-correlated activity between pairs of neurons in auditory cortex during presentation of a given tone was highly dependent on the presentation order of the tone frequencies in sequences of three tones. In the task-naive rhesus monkeys in the present study, we found some facilitated responses to sequences of three tones. In the task-naive rhesus monkeys in the present study, we found some facilitated responses to sequences of three tones. In the task-naive rhesus monkeys in the present study, we found some facilitated responses to sequences of three tones. In the task-naive rhesus monkeys in the present study, we found some facilitated responses to sequences of three tones. In the task-naive rhesus monkeys in the present study, we found some facilitated responses to sequences of three tones. In the task-naive rhesus monkeys in the present study, we found some facilitated responses to sequences of three tones. In the task-naive rhesus monkeys in the present study, we found some facilitated responses to sequences of three tones. In the task-naive rhesus monkeys in the present study, we found some facilitated responses to sequences of three tones.

Thus it is possible that many, if not most, neurons in the primary auditory cortex are sequence sensitive when stimulated with the appropriate sound sequence. In interpreting the results of our experiments, the key question is whether the response enhancement that we observed to the second, third, and fourth tones in the four-tone target melody S+ sequences are the result of learning and/or attention, or whether the enhancement can be more parsimoniously explained as the result of purely automatic, preattentive mechanisms enhancing tonal responses in the context of a tonal sequence (sensory facilitation)? If it were the latter, similar facilitation should occur in untrained monkeys. To discriminate between these two possible explanations of our data (sensory vs. learned and/or attention-driven enhancement), we compared: 1) S+ melody enhancement in the trained versus task-naive animals and 2) S+ melody enhancement in the trained animals versus C-sequence facilitation in task-naive animals. The data support differences between trained and untrained animals, thus ruling out purely sensory facilitation as the explanation (see Supplemental Fig. S2). These results therefore support facilitation due either to learning (by rerepresenting the behaviorally important tone sequence) or/and by attention. In these experiments, the data are not sufficient to distinguish between the contributions of learning and attention to the observed enhancement.

To compare the characteristics of responses to the S+ melody in trained and naive monkeys’ auditory core areas, we sought to match the distribution of cell tuning in the two populations of neurons from primary auditory cortex. As shown (Supplemental Fig. S2, A and B) the BF distribution of these two populations overlaps for over five octaves. Moreover, facilitatory and suppressive effects are largely independent of neuronal best frequency (Supplemental Fig. S2C).

A methodological limitation of our data set is the lack of physiological responses to S− tones precisely matched to the component notes in the S+ melody. Although our interpolation techniques, based on responses to near-frequency–matched single S− tones, allowed us to estimate a predicted response to the component notes of the S+ melody, it might not be possible to predict the responses to the third and fourth tones without taking into consideration the offset responses of all the previous tones, rather than just the immediately preceding tone. In our interpolation procedure, the predicted response to a component tone in the target melody is based on the sum of the offset of the just previous tone and the on-response of the current tone (derived from responses to the component tones presented alone). Although our justification for doing so...
is that we did not observe off-responses to single tones that persisted for >100 ms, more subtle effects of the previous sequence of notes may have persisted beyond that limit.

Despite this methodological limitation, our data support the presence of learned or attention-driven response enhancement in the auditory core. First, the percentage of neurons with statistically significant neuronal facilitation in the trained animals to either of the two melodies was eightfold greater than that in the naive animals (22 vs. 3%); second, there was a very different pattern in the time course of enhancement effects in the trained and task-naive monkeys. Although we did observe substantial facilitatory effects to the C-sequences in the awake, naive monkeys, these effects peaked at 75 ms after the onset of the second tone. These results are consistent with facilitation found in anesthetized monkeys (Brosch et al. 1999). In the trained animals, however, we found that facilitatory effects peaked much earlier, at 25 ms (see Fig. 9A, middle). In addition to these temporal differences in the pattern of facilitation, the amplitude of facilitatory effects was also greater in the trained animals (see Fig. 9A, top). One important caveat is that a study in the anesthetized cat A1 (Brosch and Schreiner 2000) found early (10–50 ms) facilitatory responses to two-tone sequences. Further studies are necessary to clarify the effect of training or behavioral salience on the timing of facilitatory responses in the awake animal.

In summary, these two combined lines of evidence support the view that the response enhancement to the target melody (S+; stimuli in the trained monkeys and the short latency of the enhancement, relative to results in the naive monkeys, are consequences of task-related attention and/or learning.

**Class I effects and tone-sequence suppressive responses in primary auditory cortex**

Depending on the choice of component tone frequency, amplitude, duration, and interstimulus interval, it is also possible to elicit suppressive responses to successive tonal stimuli in the core areas A1/R of auditory cortex. Neurons in A1 of the awake rabbit, anesthetized cat, and anesthetized and also awake monkey have all previously been shown to exhibit forward suppression of some two-tone sequences (Abeles and Goldstein 1972; Bartlett and Wang 2005; Brosch and Schreiner 1997, 2000; Calford and Semple 1995; Fitzpatrick et al. 1999).

In parallel with our earlier discussion of response enhancement, there are multiple lines of evidence suggesting that the suppression is strongly influenced by training. First, we found approximately twice the likelihood of statistically significant neuronal suppression to the tone sequences in either of the two melodies in the trained animals compared with the naive animals (67 vs. 36%). Second, there was also a significant difference in the peak latency of suppressive effects in the trained and task-naive monkeys. Although we did observe suppressive effects in the naive animals similar to those shown by Brosch and Schreiner (1997, 2000), these effects peaked at 25 ms after the onset of the second tone. In contrast, we found that suppressive effects peaked about 10 ms later, at 35 ms, in the trained animals. In addition to these temporal differences in the pattern of suppression, the amplitude of suppressive effects was also greater in the trained animals. Hence, as in the case of facilitation, these data suggest that many of the increased effects to the S+ melody in the trained monkeys were the result of task-related attention or learning.

**Possible behavioral strategies of monkeys performing discrimination task**

Prior to discussing the implications of these results for representation of tone sequences in auditory cortex, one may ask whether the trained monkeys actually learned the complete S+ melodies or, instead, learned to solve the task with a simpler strategy by attending to one or more features of the complex spectrottemporal pattern of those melodies. In fact, the behavioral data (Fig. 4) suggest that monkey B1, in particular, did not always need to listen to the full duration of the S+ target to perform the task successfully. Since both monkeys were initially trained to discriminate between a single S+ melody and numerous S− sounds consisting of broadband stimuli (monkey calls, FM sweeps, and broadband noise) a simple and effective strategy would be to respond to any sequence of changing narrow-band stimuli. Later in training, additional S− stimuli were added, including single pure-tone transients separated by silent gaps >1.5 s. Again, a slightly modified strategy would be to respond to any rapidly changing sequence of narrow-band stimuli presented without silent gaps. If the monkeys adopted such an implicit strategy, this might explain how monkey B2 quickly transferred its learning of the low-frequency melody, S+(L), to the high-frequency melody, S+(H). Another possible discrimination strategy would be to respond to any narrow-band tonal stimulus centered near one of the frequencies of the four notes in the S+ melody. Our behavioral data on performance errors suggest that this was indeed a prevalent, though not the exclusive, strategy used by monkey B1. In comparison, this spectral strategy does not appear to have been used extensively by monkey B2. In any case, to train the monkeys to attend carefully to the entire melody, it would have been necessary to redesign the task as a discrimination between two highly similar melodies.

In relation to the spectral strategy, one may ask how accurately monkeys can discriminate single-tone frequencies. Previous psychophysical studies have shown that, over a broad frequency range (500 Hz to 12 kHz), macaque monkeys can successfully discriminate two tones that differ in frequency by only 1–5% (Stebbins and Moody 1994). In our task, the differences in frequency between the nearest pure tone in the S− sound set and the first tone in the S+ melody were 5% for S+(L) (i.e., 504 vs. 528 Hz) and 14% for S+(H) (i.e., 5,000 vs. 5,800 Hz). Thus it is likely that by relying on pitch alone, our monkeys could have discriminated between the nearest S− sound and the first note of the target melody. However, these frequency-difference limens were measured in the context of a two-tone discrimination task, whereas our monkeys were trained to discriminate single S− tones from a multiple tone sequence. Thus although overall performance level was high for both monkeys for all S− sound sets other than pure-tone transients, both monkeys made systematic frequency-selective go-response errors. Specifically, whereas their overall performance level to single S− tones was >93% correct, they dropped to about 78% correct for the six non-target (S−) stimuli whose frequencies most closely matched the corresponding S+ tones; and monkey B1 dropped to about 50% correct for the two S− tones whose frequencies most closely
matched the first two notes of the S+(L) melody. The cost for incorrect go responses (a 3-s timeout) may have been insufficient to motivate the monkeys to reduce the number of such errors. However, if the monkeys had been using a purely spectral strategy, their performance would likely have been even lower for the closest matching S− tones, suggesting that they were probably responding at least in part to the distinctive temporal features of the S+ melodies.

Plasticity and order-sensitive representation in the auditory core

Using a paradigm to selectively enhance cortical plasticity [associative pairing of complex stimuli with nucleus basalis stimulation developed by Bakin and Weinberger (1996)], Kilgard and Merzenich (2002) explored order-sensitive changes in cortical responses. They found changes in neural responses in the auditory cortex of rats that had been entrained by associative pairing with a three-stimulus acoustic sequence, consisting of two tones (separated by 1.26 octaves) followed by a white noise burst (T1/T2/WN). They found enhanced response not only to the whole ordered three-element sequence and to both ordered two-element components (T1/T2 and T2/WN) but also to T1/N (the ordered sequence with one component missing). These and other induced changes were largely consistent with selective context-dependent facilitation for the component two-stimulus sequences contained within the full three-stimulus sequence. Our results, suggesting context-dependent response enhancement of behaviorally meaningful tone sequences following training, are in accord with their results. We speculate that, in our trained monkeys, such an order-sensitive representation of the S+ melody was elaborated, possibly based on selection and linkage of intrinsic combination-sensitive responses, and that this was coupled with enhanced responses to the individual tonal frequencies in the target melodies.

The responses observed in the trained monkeys suggest that there are multiple mechanisms to encode tonal sequences in auditory cortex. There are likely to be several levels of representation of the two S+ target melodies, including single notes, two-note transitions, multiple note phrases or sequences in a single melody, and multiple melodies with a common behavioral significance. Based on the behavioral data and the spectral strategy used by the monkeys, it is likely that there was an enhanced representation of the S+ tonal frequencies in the auditory core. However, the highest level of representation measured directly in this study was the two-note transition, in which one or more of the two-note component sequences in the four-note target melody were encoded by sequence-selective enhancement or facilitation that would not be predicted from responses to these two tones presented independently, or to the two-tone stimulus combinations (as measured in naive animals). A concatenation of two-note sequences leads to a representation of a three-note phrase or the entire four-note sequence for either target melody. We conjecture that groups of neurons could form a linked representation of the melody by combining the simple “building blocks” of responses to one-note and two-note consecutive components or into representations of “phrase pairs” (Rose et al. 2004). An even higher level is the representation of both melodies [i.e., sequence facilitation or suppression for both target melodies S+(L) and S+(H)]. We did find some cells in A1/R that showed similar responses to both melodies, with 20/63 cells showing facilitation to both and 145/186 cells showing suppression to both (Fig. 8, C and D). This suggests that neurons encoding such higher-level information about spectrotemporal pattern may reside in the core auditory areas. However, an abstract, global level of representation of the melody “gestalt” (Wright et al. 2000) probably does not occur in the core but is more likely to be found in higher levels of the auditory cortical processing stream (as suggested by physiological evidence for encoding of bird songs: Gentner and Margoliash 2003; Janata and Margoliash 1999; Margoliash and Fortune 1992; and by brain imaging studies of human melody perception: Brown et al. 2004; Halpern and Zatorre 1999; Janata 2005).

Bar-release neurons and other task-related, nonauditory activity

The responses of Class II neurons in A1 and R could not be explained by their frequency-tuning curves (generated by analyzing responses to single tones) or by any other acoustic features. Within these Class II neuronal responses, we were able to identify some that appeared to be response-related and others, reward-related. The response-related discharges best aligned to bar release and could precede as well as follow the movement. Those that preceded bar release could be interpreted as anticipatory, related to motor preparation of hand and arm movements leading to bar release. Such activity was particularly noticeable in monkey B1, trained on the S+ trial type 1 (Fig. 12), which may be related to the fact that, in this type of behavioral trial, the bar release could occur during the S+ melody and would lead directly and immediately to liquid reward. Anticipatory reward-related responses could be most clearly inferred in S+ trial type 2 (Fig. 13), in which the motor response was delayed until after the S+ melody was completed and thus the monkey could anticipate reward before releasing the bar.

If some of these neural responses in A1 and R were somatomotor in origin, they should have been present during bar grasp as well as during bar release. A dual response, to both bar grasp and bar release, was described by Brosch and colleagues, who recorded from monkeys that had been trained to categorize the direction (rising, falling) of frequency change in a tone sequence (Brosch et al. 2004, 2005; Selezneva et al. 2006). During performance, they reported that 62% of multiunits and 35% of single units responded to bar grasp and 85% of multiunits and 60% of single units responded to correct bar release. However, in our study, although we found single-unit responses aligned to the correct bar release, they constituted a much smaller percentage of total neurons (12%). Moreover, in our experiments we did not observe any neural responses in the core auditory cortical fields related to bar grasp. This discrepancy is puzzling and we can only speculate as to its explanation. It may have arisen either from: 1) different behavioral design and hence the differential salience of bar grasp and bar release in the two tasks; 2) different time course and extent of training, inasmuch as our monkeys learned the simpler S+ melody detection task within 6 mo, whereas the monkeys studied by Brosch and colleagues (2004, 2005) were trained over a period of several years in the acquisition of the more complex tone-sequence direction task; 3) different
recording techniques, i.e., perhaps we would have observed bar-grasp-related responses if we had used multiunit or local field potential responses; 4) different recording sites because, based on our extensive tonotopic mapping in both trained monkeys, all of our recordings were from A1 or R, whereas Brosch et al. (2005) indicate that they may have been recording from posterior auditory fields such as the caudomedial field (CM, which have been shown to receive a substantial somatosensory input; Hackett et al. 2007) in addition to input from the core areas of A1 and R; 5) slightly different data analyses and criteria for demonstrating correlation between spikes and bar-grasp or bar-release events; and 6) a final possibility relates to the stereotyped behavior of both monkeys during bar release that made it difficult to isolate the neuronal response to the following bar-grasp event. We measured the average time from bar release to bar regrasp and found that it averaged only 200–300 ms. Consequently, bar grasp occurred in the middle of the reward period and so the neuronal response to it may have been difficult to disentangle from the neuronal response to reward.

Clearly, further studies are needed to clarify the functional role of the neural representation of nonacoustic events in A1 and R (Ghazanfar and Schroeder 2006; Lakatos et al. 2007; Schroeder and Foxe 2005) and to understand the neural algorithms for multiplexing (and distinguishing) acoustic and nonacoustic event activity in auditory cortex.

We have shown that there is a neural representation of nonacoustic, task-related events in the auditory core that can provide a set of neural “markers” for the important events associated with the behavioral task (such as bar release and reward) that accompany the relevant acoustic events of the task or, more generally, for co-occurring environmental events (Bizley and King 2008; Kayser et al. 2008). Such unified multisensory and multiplexed responses in auditory core may be particularly valuable in the context of performing an auditory task. Both visual and somatosensory responses have been observed in auditory cortical neurons while a monkey performed an auditory task, whereas—in striking contrast—neither visual nor somatosensory activation to the same sensory stimuli was observed in the same auditory cortical cells when the monkey performed a visual-detection task (Brosch et al. 2005). This suggests that the auditory cortex is more likely to encode nonacoustic stimuli if they are meaningful for an auditory task. Thus as early as the auditory core, there may be an auditory-gated “behavioral soundtrack” that includes relevant behavioral event markers (such as the light flash signaling trial onset or bar release for reward) in synchrony with the acoustic events of the auditory task, thus forming a unified multisensory and multivalent neural representation of the behavioral task, in what is commonly thought to be a purely unisensory or modality-specific area.

Although a primary function of the core areas of auditory cortex, A1 and R, is to faithfully represent and encode incoming acoustic stimuli, recent studies suggest that core area neurons can also adjust their responses to reflect stimulus salience, allocation of attention and behavioral context (Fritz et al. 2003, 2005a,b, 2007a,b), and to encode categorical meaning (Selezneva et al. 2006). A new perspective on the function of auditory cortex is emerging, in which a major role may be to combine and coordinate multiple acoustic, multisensory, motor, and limbic inputs (Blake et al. 2002; Brosch et al. 2005; Mirabella et al. 2007; Selezneva et al. 2006; Shuler and Bear 2006; Winer et al. 2005, 2007). Our present results, showing behavioral modulation of both acoustic and nonacoustic activity in the auditory core, are consistent with this view.

ACKNOWLEDGMENTS

We thank R. Saunders for assistance with the implant surgeries on the two trained monkeys; C. Dam, H. Tak, and K. Morehead for help with animal care, behavioral training, and physiological recording at the National Institutes of Health; K. Beeman for software development and the integration of SIGNAL and National Institute of Mental Health (NIMH) CORTEX programs; G. Bertini for assistance with NIMH CORTEX; B. Tian for valuable help in developing the recording system and for assistance in initial behavioral training; M. Hauser for permission to use personal recordings of natural rhesus monkey calls from Cayo. Santiago; and M. Brosch, N. Ulanovsky, B. Scott, S. David, and S. Shamma for valuable comments and discussions. M. Mishkin designed the experiment; P. B. Yin and J. B. Fritz performed the experiments; P. B. Yin analyzed the data and prepared the figures; J. B. Fritz wrote the manuscript; and M. Mishkin, M. Sutter, and P. B. Yin helped edit the manuscript.

Present address of P. B. Yin and J. B. Fritz: The Center for Acoustic and Auditory Research, Institute for Systems Research, University of Maryland, College Park, MD 20742.

GRANTS

Much of this research was conducted while P. B. Yin and J. B. Fritz were supported by postdoctoral fellowships from the Intramural Research Program, National Institute of Mental Health, National Institutes of Health, and it was continued and completed with the assistance of additional separate grant support to J. B. Fritz, to M. Sutter, and to S. Shamma from the National Institute on Deafness and Other Communication Disorders (NIDCD).

REFERENCES


