Sensitivity of Auditory Cortical Neurons to the Locations of Leading and Lagging Sounds

Brian J. Mickey and John C. Middlebrooks
Kresge Hearing Research Institute, University of Michigan, Ann Arbor, Michigan

Submitted 27 June 2004; accepted in final form 2 April 2005

The precedence effect often is studied by presenting 2 brief sounds from different directions with a variable interstimulus delay (ISD) (reviewed by Litovsky et al. 1999). When the ISD is set below the echo threshold (typically about 5 ms), the pair of stimuli is perceived as a single fused sound. This failure to hear the lagging stimulus as a separate sound is called echo suppression. At these short ISDs, the location of the leading sound source dominates the perceived location of the sound—an observation called localization dominance. As the ISD is increased just above the echo threshold, the lagging sound is heard separately, but spatial attributes of the lagging stimulus are poorly discriminated. This reduced spatial sensitivity for the lagging sound, called lag localization suppression, is strongest at ISDs 5–10 ms (depending on the listener, type of stimulus, and psychophysical task). At ISDs 10 ms, the lagging sound is heard separately and is easily localized.

Lesion studies and physiological recordings indicate that the auditory cortex is important for sound localization and for the precedence effect. Lesions of the auditory cortex disrupt normal sound localization behavior in humans and cats (Jenkins and Masterton 1982; Jenkins and Merzenich 1984; Zatorre and Penhune 2001). Studies in humans and cats also suggest that at least one aspect of the precedence effect—localization dominance—depends on an intact auditory cortex (Cornelisse and Kelly 1987; Cranford and Oberholtzer 1976). Neurons in the cat’s auditory cortex are known to be sensitive to the locations of single sound sources (reviewed by Middlebrooks et al. 2001).

Recent studies have begun to examine the cortical representation of sounds that elicit the precedence effect. Recording from field A1 of awake rabbits, Fitzpatrick and colleagues found that the responses of most units to lagging sounds were strongly suppressed at ISDs of 1–20 ms, although a minority of units did show discrete lagging responses at delays <20 ms (Fitzpatrick et al. 1999). In a study of field A1 of anesthetized cats, Reale and Brugge (2000) only rarely observed discrete lagging responses at ISDs below 20 ms, but they did find a subset of units whose spike rates were sensitive to the presence of the lagging stimulus. In a study of fields A1 and A2 of anesthetized cats that focused on ISDs below echo threshold, we found that cortical units generally responded to paired sounds with a single burst of spikes, and that the locations signaled by spike patterns generally agreed with localization dominance and summing localization reported by listeners (Mickey and Middlebrooks 2001).

INTRODUCTION

Localization of sound is an important function of the mammalian auditory system. To represent the location of a sound source, localization cues must be extracted from signals that arrive at the 2 ears and then integrated to generate a percept or to guide a behavioral response. Human and feline listeners localize broadband sounds accurately when single isolated sounds are presented (May and Huang 1996; Middlebrooks and Green 1991; Populin and Yin 1998). In contrast, when multiple sounds arrive at the ears in quick succession, characteristic changes in localization behavior are observed in humans (Haas 1972; Wallach et al. 1949), cats (Cranford 1982; Tollin and Yin 2003), and other animals (Keller and Takahashi 1996; Kelly 1974; Spitzer et al. 2003; Wytenbach and Hoy 1993). Those changes are collectively referred to as “the precedence effect.” The auditory processing that underlies the precedence effect is thought to improve localization of sound sources and segregation of speech in commonly encountered reflective environments that produce both direct and indirect sounds (Freyman et al. 2001; Zurek 1980).
In the current study, we addressed several outstanding questions about the cortical representation of sounds that elicit the precedence effect. The long-lasting suppression of lagging responses observed in field A1 appears to conflict with observations of the precedence effect. We therefore investigated whether this suppression depends on the cortical field, the locations of leading and lagging sounds, the sound level, or the use of general anesthesia. Recordings were obtained from 3 distinct cortical fields [A1, A2, and posterior auditory field (PAF)] of anesthetized cats and from field A1 of awake cats. A range of sound levels and several combinations of leading and lagging locations were tested. We found that, regardless of these factors, discrete lagging responses were uncommon at ISDs of 10–20 ms, at which listeners report a distinct lagging sound. Even in the absence of a discrete lagging response, however, we found that the spike patterns of many cortical neurons were influenced by lagging stimuli. Finally, we tested the prediction—based on psychophysical measurements of lag localization suppression—that neurons would transmit information about the leading location regardless of the ISD, but that they would transmit information about the lagging location only as the ISD is increased ≥10 ms. In agreement with the prediction, transmission of information about the leading location was high across ISDs. With the exception of a few notable units, however, information transmission about the lagging location failed to rise as the ISD increased to as great as 16 ms.

METHODS

Animal preparations

This report describes data from 14 anesthetized cats and 5 unanesthetized cats. All procedures complied with guidelines of the University of Michigan Committee on Use and Care of Animals. All animals were purpose-bred young adults (Harlan, Indianapolis, IN). Cats studied with anesthesia were males with body weights ranging from 3.5 to 5.5 kg. The anesthetized preparation was previously detailed (Middlebrooks et al. 1998) and is only summarized here. Isoflurane anesthesia was used during surgery and intravenous α-chloralose anesthesia was used during unit recording. A craniotomy about 1 cm in diameter exposed the middle ectosylvian gyrus of the right hemisphere. A plastic retainer was cemented to the ventral margin of the opening to create a recording chamber. The animal was positioned with its head in the center of the sound chamber, its body supported in a sling with a heating pad, and its head supported by a bar attached to a skull fixture. A recording probe was inserted into active layers of the cortex. The laminar locations of recorded neurons were unknown, but we presume that most neurons were pyramidal cells in the middle layers. Probes were silicon-substrate 16-channel electrodes (Anderson et al. 1989) with inter-site spacing of 100 or 150 μm, designed for acute extracellular recording, and provided by the University of Michigan Center for Neural Communication Technology. After completely studying a cortical location, the probe was moved so that a new set of cortical units could be studied.

Cats studied without anesthesia were females with body weights ranging from 2.9 to 5.2 kg. The awake preparation was previously detailed (Mickey and Middlebrooks 2003) and is only summarized here. Each cat was trained to discriminate a 200-Hz click train (which was associated with a food reward) from noise bursts or clicks presented from random azimuthal locations. Then, using aseptic techniques under isoflurane anesthesia, a craniotomy about 1 cm in diameter was created over the middle ectosylvian gyrus of the right hemisphere and a metallic head fixture was implanted over the craniotomy. After recovery from surgery, a recording probe was placed in the auditory cortex with an external connector housed within the head fixture. Probes were silicon-substrate 16-channel electrodes (Anderson et al. 1989) with intersite spacing of 100 or 150 μm, designed for chronic extracellular recording, and provided by the University of Michigan Center for Neural Communication Technology. In the following days, unit activity was recorded in once- or twice-daily recording sessions while the cat performed the behavioral task or while it sat idly. Responses did not depend strongly on the animal’s behavior (performing vs. idle), as described previously (Mickey and Middlebrooks 2003). The cat’s head was unrestrained, so the position and orientation of the head was recorded with an electromagnetic tracking system (Polhemus fastrak, Colchester, VT). All stimulus locations are reported in head-centered coordinates (Mickey and Middlebrooks 2001). Figure 3E shows the distribution of head orientation for data described in the current report; the central 90% of measurements of head azimuth, elevation, and roll fell within the ±45° × ±5° × ±2° ranges.

After a cortical location was completely studied or responses deteriorated (typically after a few days to a few weeks), the probe was removed and a new probe was placed so that a new set of units could be studied.

Stimulus generation and delivery

Sounds were presented under free-field conditions with an apparatus that was described previously (Middlebrooks et al. 1998). In brief, a sound-attenuating chamber (dimensions 2.6 × 2.6 × 2.5 m) was lined with sound-absorbing foam to suppress reflections. Fifteen 2-way coaxial loudspeakers were positioned on a horizontal circular hoop from −60 to +60° in 10° steps, and at ±80°. Loudspeakers were located 1.2 m from the cat’s head at various azimuths. The location directly ahead of the animal was assigned an azimuth of 0°, negative azimuths were to the left (contralateral to the recorded hemisphere), and positive azimuths were to the right. Experiments were controlled by custom MATLAB software (The MathWorks, Natick, MA) running on a Pentium-based personal computer with signal processing instruments from Tucker-Davis Technologies (Gainesville, FL). Computer-controlled 2-channel D/A converters and multiplexers allowed sounds to be presented from single loudspeakers or from pairs of loudspeakers simultaneously.

Physiological experiments used clicks, noise bursts, pure-tone bursts, and click trains. The stimulus passband was 0.5–30 or 0.2–30 kHz. Because transfer functions generally differed between loudspeakers, each loudspeaker was individually calibrated by obtaining an impulse response (Zhou et al. 1992). Stimuli were created by convolution of the intended signal with the inverse impulse response of the intended loudspeaker. Click stimuli were 100-μs rectangular pulses. Noise bursts were samples of Gaussian noise with abrupt onsets and offsets; a different random sample was presented on each trial. Pure-tone bursts were ramped on and off with 5-ms raised-cosine functions. Stimulus waveforms were generated with 16-bit precision at a sampling rate of 100 kHz.

Sounds were delivered every 1–2 s in pseudorandom order such that all stimulus conditions were tested once before repeating all stimuli again in a different pseudorandom order. During experiments with anesthetized animals, we typically presented a block of 10–20 repetitions of each single-stimulus condition and a block of 10 repetitions of each paired-stimulus condition, and then repeated each block at least once more. The blocks were interleaved to reduce the potential effects of any variation of neuronal responsiveness during the 2- to 4-h stimulus set. Stimuli were presented at 2 or 3 levels in steps of 10 dB, about 20–40 dB above unit threshold. Reported sound levels refer to the sound field at the location of the cat’s head; the sound levels at the 2 tympanic membranes generally differed because of directional filtering by the pinnae. Unlike in our previous report...
Data acquisition and spike sorting

During recording sessions, the activity at 16 probe sites was recorded simultaneously and written to the computer hard disk as previously described (Mickey and Middlebrooks 2003, 2001). Spike times were expressed relative to the onset of D/A conversion, so latencies include 3.5 ms of acoustic travel time. Spike sorting was performed off-line using custom software based on principal-component analysis of spike shape, as described previously (Furukawa et al. 2000; Mickey and Middlebrooks 2003). The quality of unit isolation was characterized based on scatterplots of the first 2 or 3 principal-component weights and on histograms of interspike intervals (Mickey and Middlebrooks 2003). Single units were characterized by discrete clustering in principal-component space and a lack of interspike intervals <1 ms. In the majority of cases, we recorded multunit clusters. In those cases, spikes either could not be ascribed reliably to a single neuron or clearly originated from 2 or more neurons. In the present study, we use the term “unit” to refer to a single unit or a multunit cluster. From anesthetized animals we recorded 65 single units and 805 multunit clusters from 65 probe placements across 14 animals. From awake animals we recorded 6 single units and 38 multunit clusters from 8 probe placements across 5 animals. Most of those units were included in a previous study that described spatial sensitivity for single noise burst stimuli (Mickey and Middlebrooks 2003).

Data analysis

The assignment of each unit to a cortical field was based on pure-tone frequency-tuning curves and the location of recording sites relative to gross anatomical landmarks. Units in field A1 were distinguished by narrow frequency tuning (half-maximal bandwidth ≤1.33 octaves at about 40 dB above threshold) and by locations on the dorsal part of the middle ectosylvian gyrus. Units in the dorsal zone of A1 (field DZ) were dorsal and caudal to A1 proper and were distinguished by complex frequency tuning (band-pass, low-pass, highpass, or multipeaked tuning curves). Units in field A2 were located ventral to A1 and exhibited complex frequency tuning. Units in the posterior auditory field (PAF) were located within or near the anterior bank of the posterior ectosylvian sulcus, they exhibited complex frequency tuning, and they typically responded with median first-spike latencies >25 ms. Of 870 units from anesthetized animals, 271 (31%) were from A1, 13 (1.5%) were from DZ, 466 (54%) were from A2, and 103 (12%) were from PAF. For 17 units (2%), the cortical field was unassigned. Units recorded from awake animals were located in A1 (Mickey and Middlebrooks 2003).

To quantify the suppression of unit responses to lagging sounds, we calculated recovery times for each unit. First, we determined the average number of spikes per trial obtained when an isolated click was delivered from the lagging loudspeaker (i.e., in the absence of a leading stimulus). This number was estimated by counting spikes in a 30-ms window centered on the time of peak response to that loudspeaker. Second, for each ISD tested with paired clicks, we determined the average number of spikes per trial in a 30-ms window centered at the expected time of a lagging response (i.e., the window was delayed by a time equal to the ISD). The normalized spike count arising from the lagging click at each ISD was computed by dividing by the number of spikes recorded after an isolated click. Third, we found the smallest tested ISD at which the normalized spike count was ≥0.10, 0.50, and 0.90. Recovery times $t_{0.10}$, $t_{0.50}$, and $t_{0.90}$ were calculated by linearly interpolating between this ISD and the next smallest tested ISD and finding the ISD at which the normalized spike count crossed 0.10, 0.50, and 0.90. In some cases, recovery times were undefined because the response to the isolated click was too small (the mean spike count across trials was <2 SEs above zero).

To estimate the amount of information transmitted by unit responses about source location or ISD, we used a pattern-recognition algorithm and information-theoretic analysis. The procedure is described elsewhere in detail (Mickey and Middlebrooks 2003) and summarized here. The pattern-recognition approach was favored because it is sensitive to stimulus-related information carried by neuronal spike times. An alternative would have been to reduce multidimensional temporal patterns to one-dimensional spike counts. We tested a spike-count-only analysis, but found that elimination of spike-time information reduced transmitted information by about 36%; similarly, our previous work has demonstrated that spike counts transmit 37–56% less information than do full spike patterns (Furukawa and Middlebrooks 2002; Mickey and Middlebrooks 2003; Middlebrooks et al. 1998). Because the issue of information loss is not central to the current experimental questions, we illustrate only the results of the analysis that used full spike patterns.

The spike patterns of each unit were analyzed in 3 steps. In the first step, we randomly divided trials into 2 sets of equal size and computed average response patterns $R_A(s)$ and $R_B(s)$ for each set, where $s$ is the set of unique combinations of stimulus parameters. Response patterns consisted of spike density functions (1-ms time bins, 1-ms Gaussian convolution) over the range of time 10–60 ms after stimulus onset for each location or ISD.

In the second step, we performed pattern recognition on $R_A$ and $R_B$ using a probabilistic neural network constructed using the MATLAB Neural Network Toolbox (Demuth and Beale 2000). Network weights and biases were calculated such that the input of each $R_A$ resulted in an output $Y_A(s)$ that corresponded to the true source location or ISD. To characterize how consistent $R_A$ was with $R_A$, $R_B$ was presented to the network and the output $Y_B(s)$ was recorded. $R_A$ and $R_B$ were then interchanged and an output $Y_A(s)$ was obtained in a similar manner. The greater the similarity of stimulus-related responses in $R_A$ and $R_B$, the more closely $Y_A(s)$ and $Y_B(s)$ estimated the true stimulus values. The entire procedure was repeated a total of 50 times (i.e., 50 different randomly derived responses, $R_A$ and $R_B$ were analyzed), yielding 100 outputs for each location or ISD.

In the third step, the amount of stimulus-related information transmitted by the pattern recognition was computed. A confusion matrix was constructed from the 100 network outputs $Y_A(s)$ and $Y_B(s)$. Element $i,j$ of the confusion matrix consisted of the number of network outputs at stimulus value $i$ for the true stimulus value $j$. The accuracy of network estimates was characterized by computing the transmitted information (T), i.e., the mutual information (Rieke et al. 1997). The more closely that network outputs estimated the true stimulus locations, the greater the value of T. In general, T was greatest when average unit responses for various stimulus locations were easily discriminable; T was near zero when responses varied little between different stimuli. Chance variation of spike patterns with location or ISD led to overestimation of T, so we ultimately used the corrected transmitted information, $T_{\text{cor}} = T - T_0$, where $T_0$ was determined by a control analysis. In the control analysis, the trial-by-trial association of stimulus location with unit responses was reassigned randomly. The subsequent 3-step analysis was identical to that used to determine T. Rather than computing $T_0$ only once as previously described (Mickey and Middlebrooks 2003), we computed $T_0$ 5–20 times and took the mean value to further reduce noise originating from randomization.

To determine whether a unit transmitted a statistically significant amount of stimulus-related information, we used a distribution-free test: if T was greater than the first 20 independently calculated values of $T_0$, then $T_{\text{cor}}$ was considered significant ($P < 0.05$). Throughout this report we use the relative information, $T_{\text{rel}} = T_{\text{cor}} / T_{\text{max}}$, where $T_{\text{max}}$ is the maximum theoretical information (i.e., the entropy of the stimulus set, $\Sigma p_i \log_2 [p_i]$, where $p_i$ is the probability of stimulus type...
The relative information reflects the total amount of information transmitted by a particular unit during the recording session.

RESULTS

We recorded 65 single units and 805 multiunit clusters from fields A1, A2, and PAF of anesthetized cats while presenting single clicks and paired clicks under free-field conditions. Clicks were delivered from loudspeakers positioned at various azimuths in the frontal hemifield (−60 to +60°) over a range of sound levels (20–50 dB above unit threshold). Paired clicks were presented, one click from each of 2 loudspeakers, with ISDs that ranged from 1 to 300 ms.

Suppression of lagging responses

Although listeners report a distinct lagging sound at ISDs ≥5 ms, unit responses typically lacked a discrete component that corresponded to the lagging sound as the ISD was increased to >5 ms. Figure 1, A–C shows a raster display for a representative unit. In response to single clicks, this unit responded strongly to contralateral clicks (−50°) and weakly to ipsilateral clicks (+50°) (Fig. 1A). Although an ipsilateral click failed to reliably evoke a response, it nonetheless produced complete suppression of the response to a contralateral click presented 1–25 ms later (Fig. 1B). This suppression was long lasting: the unit responded to the lagging sound only when the ISD was ≥150 ms (Fig. 1C). Figure 1, D–F shows one of the few units in our population that responded strongly and discretely to the lagging click at ISDs <25 ms (Fig. 1E). We classified units as showing a discrete lagging response if the number of spikes elicited within the expected range of latency was significantly greater when a lagging click was present than when it was absent (P < 0.01, Wilcoxon rank-sum test). Only 34 of 290 units (12%) showed a discrete lagging response at an ISD of 20 ms. At a delay at which human and cat listeners report a distinct lagging sound. Thus for the vast majority of units, the suppression that followed the leading sound persisted beyond ISDs at which listeners report hearing the lagging sound.

To quantify the suppression of lagging responses across our unit population, we calculated recovery times $t_{10}$, $t_{50}$, and $t_{90}$ for each unit from responses to paired clicks like those shown in Fig. 1. These recovery times estimated the smallest ISD at which the magnitude of the lagging response was 10, 50, and 90% (respectively) of the response to an isolated click from the lagging location. Loudspeakers were positioned at 3 locations: contralateral to the recording site (−40 to −50°), midline (0°), and ipsilateral (+40 to +50°). Three different lead–lag loudspeaker configurations were tested: midline–contralateral, ipsilateral–contralateral, and contralateral–ipsilateral. Figure 2 shows cumulative distributions of recovery times for units in 3 cortical fields with various loudspeaker configurations. Each curve plots the fraction of units that had recovered at a given ISD; black curves represent $t_{50}$ and the top and bottom gray curves represent $t_{10}$ and $t_{90}$, respectively. As shown in Fig. 2, A–C, suppression lasted longest for units in PAF. At an ISD of 300 ms, the proportion of units that had recovered to at least the half-maximal value was 32% in PAF, 47% in A1, and 60% in A2 (median $t_{50}$ >300 ms in A1, 249 ms in A2, >300 ms in PAF). Furthermore, comparisons of Fig. 2, A–B, D–E, and F–G suggest that suppression of lagging responses was stronger when the leading stimulus was from a midline or contralateral location (i.e., locations that elicited a greater response). In particular, the ipsilateral–contralateral configuration produced the least suppression, with practically all units recovering half-maximally by 300 ms (median $t_{50}$, 116 ms in A1, 151 ms in A2). Each pairwise comparison between different cortical fields and between different loudspeaker configurations was statistically significant ($P < 0.05$, Wilcoxon rank-sum test). Regardless of cortical field or loudspeaker configuration, a large majority of units exhibited long-lasting suppression of lagging responses.

Because spatial receptive fields are known to be level dependent in anesthetized preparations (Imig et al. 1990; Middlebrooks and Pettigrew 1981; Middlebrooks et al. 1998; Rajan et al. 1990; Reale et al. 2003; Stecker et al. 2003), we wondered whether suppression of lagging responses might depend on the sound level. We addressed this question by comparing recovery times obtained at a sound level 40–50 dB above unit threshold to those obtained at a sound level 20 dB below that. The levels of the 2 clicks were always equal. Among the 7 combinations of cortical field and loudspeaker configuration that were tested, recovery times at the higher sound level were longer in 2 conditions and shorter in one condition ($P < 0.05$,
Wilcoxon sign-rank test). Nonetheless, the magnitude of the change in recovery times with sound level was modest (<25 ms) in each case. For example, for the ipsilateral–contralateral configuration, a 20-dB increase in sound level increased the median $t_{50}$ from 112 to 116 ms in field A1 ($n = 34$, $P = 0.15$) and from 144 to 153 ms in field A2 ($n = 64$, $P = 0.01$). Thus suppression of lagging responses did not depend strongly on stimulus intensity.

Responses in awake animals

The findings of previous studies conducted with or without anesthesia suggest that the use of general anesthesia may augment the suppression of lagging responses (Fitzpatrick et al. 1999; Reale and Brugge 2000). We addressed this question by recording from field A1 using an awake cat preparation (Mickey and Middlebrooks 2003). Forty-four click-responsive units were recorded from 5 animals. Figure 3 shows peristimulus time histograms for the ipsilateral–contralateral configuration for 3 representative units. The unit shown in Fig. 3A was unusual among our population in that it exhibited clear lagging responses even at ISDs <20 ms. Note that this unit showed little or no response to the leading click (best illustrated at an ISD of 100 ms). The example in Fig. 3B was more typical in that this unit showed discrete lagging responses at ISDs $\geq$50 ms, but not at ISDs $<20$ ms.

As with units in anesthetized cats, we classified units as showing a lagging response if the number of spikes elicited by paired clicks within the expected range of latencies was significantly greater than the background firing rate ($P < 0.01$, Wilcoxon rank-sum test). Figure 3D compares responses in the awake preparation to those in the anesthetized preparation for the ipsilateral–contralateral loudspeaker configuration. Filled circles represent data from awake animals (11–41 units from 2–7 probe placements in 1–4 animals, depending on the ISD tested). Open circles represent comparable data from anesthetized animals (290–446 units from 20–31 probe placements in 7–9 animals, depending on the ISD tested). It should be noted that there was a significant amount of heterogeneity between probe placements and between animals in the anesthetized preparation ($P < 0.05$, chi-square test for homogeneity among dichotomous populations). Nonetheless, the overall proportion of units that showed a lagging response at a given ISD was greater in awake animals than in anesthetized animals. Thus suppression was not as long-lasting in awake animals. Because spontaneous rates also tended to be greater in the awake preparation, we sometimes observed suppressive responses. Figure 3C shows a unit that responded to leading and lagging clicks with a decrease in firing below the baseline rate. In summary, suppression was shorter lasting in awake animals, but the great majority of units nonetheless failed to respond to lagging sounds at ISDs above behavioral echo thresholds.

Modulation of spike patterns by lagging sounds

As described above, the vast majority of units failed to exhibit discrete responses to lagging sounds at ISDs <20 ms. Nonetheless, many of these units exhibited spike rates or spike timing that depended on the ISD. That is, the response to the leading sound was modulated by the presence of a lagging sound. Two examples are shown in Fig. 4. The unit in Fig. 4, A–C mainly exhibited variation in spike timing (Fig. 4B), but also showed some variation in spike rate (Fig. 4C), as a function of ISD. The unit in Fig. 4, D–F showed some variation in spike timing (Fig. 4E) and a strong dependency of spike rate (Fig. 4F). Units such as these indicated that, in contrast to the few units that explicitly represented the ISD with a discrete lagging response (for example, Fig. 1, D–F), many units might represent lagging sounds implicitly by way of their spike rates or spike patterns. We reasoned that, if units show implicit sensitivity to lagging sounds, then they should be considered part of the cortical representation of paired sounds.

To investigate this idea, we quantified the sensitivity of each unit to ISD for ISDs in the range 1–20 ms. Unit responses were recorded from anesthetized animals for 6–8 different ISDs ranging from 1 to 16–20 ms. Responses were then analyzed using a pattern recognition algorithm and information-theoretic analysis (see Data analysis in METHODS). This analysis esti-
mated the amount of information transmitted by spike patterns about the ISD of the stimulus. The transmitted information ranged from 0 to nearly 2 bits across our unit population. To account for different numbers of ISDs tested among different units, for each unit we calculated the relative information (range 0 –1), which was the actual transmitted information divided by the theoretical maximum transmitted information [i.e., the entropy of the stimulus set, log2 (n) bits, where n is the number of ISDs tested].

Figure 5 shows transmitted information for our unit population for various loudspeaker configurations. The greatest information transmission was noted when the leading sound was ipsilateral and the lagging sound was contralateral to the recorded hemisphere (ipsilateral–contralateral configuration, Fig. 5A). For this configuration, the mean relative information across the population was 0.08, with 40% of units transmitting an amount of information about ISD that was statistically significant (P < 0.05, gray bars in Fig. 5). In contrast, the contralateral–ipsilateral configuration (Fig. 5B) resulted in low amounts of transmitted information (mean relative information 0.01, 10% significant). Similar to the above analysis of recovery times (Suppression of lagging responses), information transmitted about the delay of the lagging stimulus was less when the leading location was more effective (compare Fig. 5, A to C, B to E, and D to F). The information transmitted by units in cortical fields A1 and A2 was not significantly different under any stimulus configuration (P > 0.05, Wilcoxon rank-sum tests). Compared with A1 and A2, information transmission by PAF units (n = 20) tended to be greater for the midline–contralateral configuration (PAF median 0.047; A1 median 0.010, P = 0.03; A2 median 0.013, P = 0.09) and less for the ipsilateral–midline configuration (PAF median −0.008; A1 median 0.014, P = 0.03; A2 median 0.042, P = 0.0004).

Transmission of information about leading and lagging locations

The previous analyses showed that many units were sensitive to the presence of a lagging sound. We were specifically interested in whether the spike patterns of units could also accurately signal the location of the lagging sound. Furthermore, for units that were sensitive to the lagging location, we...
wanted to know how that sensitivity depended on the ISD. To address those questions, we presented paired sounds under 2 conditions: with a variable leading location and fixed lagging location; or with a fixed leading location and variable lagging location. In each case, the fixed location was always 0° azimuth and the variable location ranged from −60° to +60° azimuth, in 20° steps. In addition, the ISD of paired sounds was varied among 6 values within the range 1–16 ms. Psychophysical experiments in humans have shown that localization of leading sounds is accurate at all ISDs, whereas localization of lagging sounds is accurate only at ISDs >5–10 ms. Therefore we predicted that under conditions in which the leading location varied, location-related information transmission would be high regardless of the ISD. In contrast, under conditions in which the lagging location varied, information transmission was predicted to be low at small ISDs (1–8 ms) and higher at large ISDs (12–16 ms).

Figure 6 shows an example of a unit whose responses agreed with those predictions. In response to single clicks, the unit responded more strongly to contralateral sounds than to ipsilateral sounds (Fig. 6A). Figure 6B shows responses when the leading location was varied and the lagging location was fixed at 0°. The unit responded with a single burst of spikes at an ISD of 1 ms (Fig. 6B, top), but a discrete lagging response emerged as the ISD was increased to 16 ms (Fig. 6B, bottom). Although the particular location-dependent temporal pattern of spikes varied with ISD, spike patterns were modulated by the leading location at every ISD tested. Notably, at ISDs of 12–16 ms, the leading location was varied and the lagging location fixed, both leading and lagging components of spike patterns showed sensitivity to the leading location (Fig. 6B, bottom 2 panels). As described above for longer ISDs (Suppression of lagging responses), this unit’s lagging responses were suppressed to a greater extent when the leading stimulus was presented from more effective locations (contralateral). Figure 6C shows responses when the leading location was fixed at 0° and the lagging location was varied. Responses appeared to depend on the lagging location when the ISD was 8–16 ms but not when the ISD was 1–4 ms (Fig. 6C). Inspection of these spike patterns suggested that only the lagging component of the response was location sensitive at ISDs of 12–16 ms (Fig. 6C, bottom 2 panels). It is notable that, under both leading-variable and lagging-variable conditions, lagging responses were diminished relative to responses to single stimuli from the same locations.

We estimated the amount of information that unit spike patterns transmitted about location using a pattern recognition algorithm and information-theoretic analysis (see Data analysis in Methods). Information transmission is shown as a function of ISD under conditions in which we varied the leading location (Fig. 6D) or the lagging location (Fig. 6E). When the leading location was varied, this unit showed relatively high transmitted information across all ISDs tested, although transmission was somewhat greater at larger ISDs (Fig. 6D). When the lagging location was varied, the unit transmitted essentially no information at the leading location at ISDs of 1–4 ms, but a significant amount of information (P < 0.05) at ISDs of 8–16 ms (Fig. 6E). Thus the responses of this unit agreed qualitatively with human psychophysical results in that sensitivity for the leading location was relatively high across ISDs whereas sensitivity for the lagging location was impaired at shorter ISDs.

Units such as this one were uncommon. Only 3 of 134 units (2%) showed a dependency of transmitted information on ISD like that shown in Fig. 6, D and E (i.e., a significant amount of information at ISDs of 12–16 ms but not 1–4 ms). More often we found that transmission of location-related information was high when the leading location varied and low when the lagging location varied, regardless of ISD. Distributions of transmitted information across our unit population are shown in Fig. 7. Among 205 units, the information transmitted about the location of single clicks ranged from about 0 to nearly 2 bits (Fig. 7A), with 129 units (63%) showing a statistically significant amount of information (P < 0.05, gray bars in Fig. 7A). [The entropy for this set of 7 stimulus locations is log2 (7), about 2.8 bits, so 2 bits is equivalent to a relative information of 0.71.] These 129 location-sensitive units were analyzed further to determine the amount of information transmitted under paired-stimulus conditions in which either the leading location (Fig. 7B) or the lagging location (Fig. 7C) was variable. When the leading location was varied, units showed relatively high transmission of location-related information: across ISDs of 1–16 ms, the median relative information was 0.10–0.18, with 55–78% of units transmitting a statistically significant amount of information (Fig. 7B). In contrast, when the lagging location was varied, units transmitted little location-related information: across ISDs of 1–16 ms, the median relative information was 0.10–0.18, with 55–78% of units transmitting a statistically significant amount of information (Fig. 7B).
relative information was 0–0.02, with only 2–23% of units transmitting a statistically significant amount of information (Fig. 7C). Figure 7, D and E shows transmitted information as a function of ISD for the population: open circles represent the mean and solid curves indicate 10th, 50th, and 90th percentiles. In summary, units transmitted a high amount of information when the leading location varied and a low amount when the lagging location varied, with only a weak dependency on the ISD.

We reached the same conclusions when this analysis was applied to mean spike rates rather than full spike patterns. The overall amount of information transmitted was lower when spike rates were used (e.g., mean transmitted information was 36% lower in the single-stimulus condition), but the dependency of transmitted information on ISD was otherwise indis-

FIG. 6. Information transmitted about leading and lagging locations for an individual unit. A: raster plot showing the responses of one unit to single clicks presented from various frontal locations. For clarity, lines separating different locations are omitted. B: raster plots of responses to paired clicks where the leading location was varied and the lagging location was fixed at 0°. Each panel represents responses to sounds at a given ISD, as indicated beside the panel. C: raster plots of responses to paired clicks where the leading location was fixed at 0° and the lagging location was varied. D: information transmitted about leading location for the responses shown in B. Relative information is plotted vs. ISD. E: information transmitted about lagging location for the responses shown in C. Relative information is plotted vs. ISD.

FIG. 7. Information transmitted about leading and lagging locations for the unit population. Relative information is represented in each panel. A: information transmitted about the location of single clicks. Distribution of information is shown for 205 units. Units that transmitted an amount of information that was statistically significant ($P < 0.05$) are shown as gray bars ($n = 129$, 63%). B: information transmitted about the location of leading clicks. Distributions of information are shown for the 129 significant units from A for various ISDs, as indicated beside each panel. Units that transmitted a significant amount of information about the leading location are represented by gray bars. C: information transmitted about the location of lagging clicks. Distributions of information are shown for the 129 significant units from A for various ISDs, as indicated beside each panel. Units that transmitted a significant amount of information about the lagging location are represented by gray bars. D: information transmitted about leading location vs. ISD. Open circles represent the mean for the population; curves represent the 10th, 50th, and 90th percentiles. E: information transmitted about lagging location vs. ISD. Open circles represent the mean for the population; curves represent the 10th, 50th, and 90th percentiles.
tinguishable from that shown in Fig. 7, D and E. That finding indicates that, in response to paired stimuli, both spike rate and spike timing carried location-related information, consistent with previous studies that used single stimuli (Furukawa and Middlebrooks 2002; Mickey and Middlebrooks 2003; Middlebrooks et al. 1998).

**D I S C U S S I O N**

In the current study, we report that the responses of most auditory cortical neurons to a lagging sound were suppressed when a leading sound preceded it by up to tens of milliseconds. This suppression was found regardless of the leading and lagging locations, the sound level, the cortical field, or the state of anesthesia. Nonetheless, at ISDs of 1–20 ms, the firing patterns of many cortical units varied with ISD, indicating that they were sensitive to the presence of a lagging sound, even when there was no discrete lagging response. Overall, units transmitted relatively high amounts of information about the location of a leading stimulus and low amounts of information about the location of a lagging stimulus, with minimal ISD dependency. Only a small fraction of units responded in a way consistent with psychophysical observations in humans or cats.

**Suppression in auditory cortex**

The suppression of lagging responses that we observed is qualitatively similar to previous reports. Fitzpatrick and colleagues (1999) presented dichotic click stimuli while recording from field A1 in awake rabbits and found that the median recovery time for their unit population was 22 ms. We found somewhat more suppression, with only 14% of units showing discrete lagging responses at ISDs of 16–20 ms in A1 of awake cats. Reale and Bruggé (2000) used virtual acoustic spatial stimuli while recording from A1 of anesthetized cats and found a median recovery time of 73 ms when the lagging sound was from the neuron’s preferred location and the leading sound was from a nonpreferred location. By comparison, we found that only 16% of units in A1 of anesthetized cats had a \( t_{lag} < 73 \) ms with the ipsilateral–contralateral configuration. These differences in recovery times may be related to the method of stimulus presentation. It is not known how echo thresholds differ when using virtual acoustic spatial stimuli, but echo thresholds are known to be smaller for dichotic stimuli than for free-field sounds (Litovsky et al. 1999). Aspects of the animal preparations such as species or type of anesthetic may also have contributed to the variation of results between studies. Nonetheless, all 3 studies agree in that the great majority of cortical neurons are suppressed at ISDs beyond behavioral echo thresholds, even in the absence of anesthesia. It will be of interest for future studies to determine whether different attentional states or different behavioral contexts result in differing amounts of suppression of lagging responses. Although we have not found a behavioral dependency (Mickey and Middlebrooks 2003), the responses of auditory cortical neurons can vary with the animal’s behavior (e.g., sound localization; Benson et al. 1981).

The paucity of lagging responses that we observed at short ISDs is consistent with previous cortical studies that used repetitive stimuli. In those studies, each repetition of the stimulus possesses the same spatial or binaural characteristics, so in terms of the precedence effect, such stimuli correspond to an unrealistic situation in which a direct sound and its echos arise from the same direction. Nonetheless, the comparison may be useful for the temporal aspects of the precedence effect. In general, for repetition rates \( \geq 50 \) Hz (periods \( \leq 20 \) ms), cortical neurons do not show entrainment (i.e., a response to every stimulus repetition), although many neurons do show phase locking at high rates (Bieser and Muller-Peuss 1996; de Ribaupierre et al. 1972; Lu and Wang 2000; Lu et al. 2001; Phillips et al. 1989; Schreiner and Urban 1986, 1988; Stein-Schneider et al. 1998). Our results also agree with previous studies that found that phase locking extends to higher repetition rates in awake animals compared with anesthetized animals (Bieser and Muller-Peuss 1996; Lu et al. 2001; Stein-Schneider et al. 1998) and in A1 compared with A2 or PAF (Schreiner and Urban 1988).

What is the source of the long-lasting suppression that we and others have observed in the auditory cortex? The failure to respond to lagging sounds is likely to be inherited, to some extent, from earlier auditory centers. The characteristics of suppression in the auditory thalamus are unknown, but neurons in the inferior colliculus show suppression similar to that in the cortex, although with a shorter duration (Fitzpatrick et al. 1999; Litovsky and Yin 1998; Yin 1994). The additional suppression found in the cortex seems likely to arise from inhibitory inputs to cortical neurons. Intracellular recordings suggest that the typically phasic responses of cortical neurons may result from an excitatory postsynaptic potential followed a few milliseconds later by a more prolonged inhibitory postsynaptic potential (de Ribaupierre et al. 1972; Wehr and Zador 2003). Whereas the suppression is likely to arise (at least in part) by inhibition, it does not appear to be spatial lateral inhibition. That conclusion is based on the finding that, as described here and in another study (Reale and Bruggé 2000), suppression of lagging clicks is generally longer lasting when the leading sound is delivered from a more effective location than when it is delivered from a less-effective location. The suppression may therefore be better characterized as adaptation. Consistent with the idea of adaptation, the suppression is long lasting regardless of whether leading and lagging sounds are presented from the same location or from different locations (Fitzpatrick et al. 1999; Reale and Bruggé 2000). A definitive answer to this issue will require measurement of postsynaptic potentials while presenting click pairs.

**Cortical representation of lagging sounds**

The results of physiological studies thus indicate that cortical neurons rarely respond to lagging sounds at ISDs near the behavioral echo thresholds of humans and cats. Presumably, the responses of cortical neurons underlie spatial aspects of the precedence effect, so the question arises: How does the cortex represent lagging sounds and their spatial attributes? We consider 2 possibilities here. If we assume an explicit representation of lagging sounds, that is, a representation based on discrete lagging responses, then we are forced to place the burden of this representation on a small minority of units (Fitzpatrick et al. 1999; Yin 1994). In other words, psychophysical performance is attributed to a more-or-less functionally distinct subpopulation of neurons. In support of this view, we found a small number of units (a few percent of our
population) that responded to lagging sounds at ISDs near the behavioral echo thresholds of humans and cats and a similar small number that showed spatial sensitivity for lagging sounds.

Alternatively, the representation of lagging sounds may be implicit. In this view, the neuron lacks a discrete lagging response, yet its spike rate or temporal pattern of spikes is sensitive to attributes of the lagging sound. In support of this view, we found that the responses of up to 40% of units showed some sensitivity to the ISD of the stimulus. This observation has a parallel in psychophysics: human listeners are able to discriminate between stimuli with different ISDs even when the ISD is below echo threshold (Blauert 1997). We found that implicit ISD sensitivity is fairly common, so the implicit representation of sounds would not require one to posit a small minority of specialized units. Similar implicit representations have been proposed in other contexts. Work from our laboratory has shown that the temporal spike patterns of cortical neurons can accurately represent the locations of single sounds (Furukawa and Middlebrooks 2002; Mickey and Middlebrooks 2003; Middlebrooks et al. 1998). An implicit rate-based code has been proposed for the representation of click trains with small interstimulus intervals in the auditory cortex (Lu et al. 2001). Of course, implicit and explicit representations of lagging sounds are not mutually exclusive: some neurons might represent lagging sounds implicitly whereas others represent them explicitly.

We measured information transmission for the locations of leading and lagging sounds with the goal of comparing the results with human psychophysical measurements. In general, localization (or discrimination between locations) of leading sounds is relatively accurate, whereas localization of lagging sounds is accurate only at longer ISDs (Litovsky et al. 1999). Thresholds for lag-discrimination suppression have been estimated at 5–10 ms (Yang and Grantham 1997). Our results agree with psychophysical results in that substantially more information was transmitted about leading sounds than about lagging sounds, but disagree in that the vast majority of units showed no significant transmission of information about the lagging location at ISDs of 8–16 ms. Our method for estimating information transmission is independent of the type of coding used by the neuron (e.g., implicit or explicit). Therefore regardless of the coding strategy, it appears that a small subpopulation of specialized neurons must be invoked to account for localization of lagging sounds.

ACKNOWLEDGMENTS

We thank E. Macpherson, C. Stecker, and S. Furukawa for help with experiments and for many useful discussions; Z. Omsan for technical assistance; and E. Macpherson and I. Harrington for valuable comments on an earlier version of the manuscript.

Present address of B. J. Mickey: Department of Psychiatry, University of Michigan, Ann Arbor, MI 48109-0118 (E-mail: bmickey@umich.edu).

GRANTS

This work was supported by National Institutes of Health Grants RO1 DC-00420, RO1 EB-00308, T32 GM-07863, T32 DC-0001, and P30 DC-00420, RO1 EB-00308, T32 GM-07863, T32 DC-0001, and P30 DC-00420, and by NIH Grant P41 RR-09754.

REFERENCES


