Directional Sensitivity of Neurons in the Primary Auditory (AI) Cortex of the Cat to Successive Sounds Ordered in Time and Space

RICHARD A. REALE AND JOHN F. BRUGGE
Department of Physiology and Waisman Center, University of Wisconsin, Madison, Wisconsin 53705
Received 8 December 1999; accepted in final form 3 April 2000


Two transient sounds, considered as a conditioner followed by a probe, were delivered successively from the same or different direction in virtual acoustic space (VAS) while recording from single neurons in primary auditory cortex (AI) of cats under general anesthesia. Typically, the response to the probe sound was progressively suppressed as the interval between the two sounds (ISI) was systematically reduced from 400 to 50 ms, and the sound-source directions were within the cell’s virtual space receptive field (VSRF). Suppression of the cell’s discharge could be accompanied by an increase in response latency. In some neurons, the joint response to two sounds delivered successively was summative or facilitative at ISIs below about 20 ms. These relationships held throughout the VSRF, including those directions on or near the cell’s acoustic axis where sounds often elicit the strongest response. The strength of suppression varied systematically with the direction of the probe sound when the ISI was fixed and the conditioning sound arrived from the cell’s acoustic axis. Consequently a VSRF defined by the response to the lagging probe sound was progressively reduced in size when ISIs were shortened from 400 to 50 ms. Although the presence of a previous sound reduced the size of the VSRF, for many of these VSRFs a systematic gradient of response latency was maintained. The maintenance of such a gradient may provide a mechanism by which directional acuity remains intact in an acoustic environment containing competing acoustic transients.

INTRODUCTION

Natural environments are filled with transient sounds that arrive at the two ears in unpredictable succession, either directly from primary sources or indirectly as a result of reflections from objects near and far. The central auditory system has devised mechanisms, distributed between and including the lower brain stem and cortex, to extract information about the identity and direction of sounds in this competitive environment when the time interval between the transients ranges from tens of microseconds to hundreds of milliseconds.

Psychophysical localization studies employing successive sounds have focused attention mainly on the precedence effect, a term that encompasses a complex set of perceptual phenomena related to sounds that are separated in time by less than about 50 ms (reviewed by Blauert 1997; Litovsky et al. 1999; Zurek 1987). These studies have usually been cast in the context of a listener’s ability to localize and identify a sound source that competes with its early reflections in a reverberant space. Neurons in the auditory brain stem, midbrain, and cortex are sensitive to temporal separations of a few tens of milliseconds and thus are considered part of the neural circuitry that underlies these phenomena (Carney and Yin 1990; Fitzpatrick et al. 1995, 1997; Litovsky and Yin 1998a,b; Mickey and Middlebrooks, 2000; Mickey et al. 1999; Yin 1994). Results of psychophysical studies of spatial resolution (Grantham 1986; Perrott and Pacheco 1989), source-echo detection (Stellmack et al. 1997), localization aftereffects (Thurlow et al. 1965), and motion discrimination (Saberi and Hafter 1997) have also made clear, however, that time constants involved in processing directional information may range from tens to hundreds of milliseconds, which is considerably longer than that required for precedence phenomena (see also Blauert 1972; Grantham and Wightman 1978; Perrott 1982). Successive sounds having these longer time intervals have also been reported to have profound influences on lateralization judgments (Hari 1995; Stellmack and Lutfi 1996; Stellmack et al. 1997). Hartmann (1997) refers to spatial perceptual phenomena with such long time constants as de-reverberation, a situation found in large concert halls in which a listener is not fully aware of reverberant sounds even though their energy may be many times that of the incident signals. All of these latter observations suggest that even though a listener may hear two sounds in space separated by time intervals ranging from tens to hundreds of milliseconds, information about each of their directions may not be processed entirely independently.

The fidelity by which a primary auditory cortex (AI) neuron responds to a probe stimulus depends on when in the past a conditioning stimulus occurred. Such data have often been presented in the form of “recovery functions,” which plot changes in response strength to the probe as a function of conditioner-probe interval. Despite methodological and species differences among studies of this kind, the range over which suppression has been shown to occur is in the hundreds of milliseconds (Altman et al. 1997; Borsanyi 1964; Brosch and Schreiner 1997; Calford and Semple 1995; Fitzpatrick et al. 1997; Howard et al. 2000; Lee et al. 1984; Liegeois-Chauvel 1991; Perl and Casby 1954; Puletti and Ceslesia 1970; Rosenblith 1950; Rosenblith and Rosenzweig 1951; Rosenzweig 1951, 1954; Rosenzweig and Rosenblith 1950, 1953; Serkov and Yanovskii 1970). Long AI recovery times have usually

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
been interpreted in terms of forward masking and lateral inhibition (e.g., Brosch and Schreiner 1997; Calford and Semple 1995) or coding of temporal intervals, such as occurs in speech and communication sounds (e.g., Creutzfeldt et al. 1980; Eggermont 1991, 1994; Phillips and Farmer 1990; Phillips and Hall 1990; Schreiner et al. 1997). These studies were typically carried out under monaural or binaural earphone listening conditions using sounds that did not include auditory spatial cues. Considering the persistent effects that a binaural conditioning sound may have on the spatial perception of a subsequent sound, we naturally questioned the extent to which directional sensitivity of an AI neuron to a directional probe, as reflected in the spatial receptive field, may likewise be affected by a conditioning stimulus arising from the same or different direction. We showed previously that AI neurons have broad spatial receptive fields when their directional selectivity is tested with single transient sounds that were synthesized in an otherwise anechoic virtual acoustic space (Brugge et al. 1994, 1996). These receptive fields typically exhibit functional gradients, and an information theoretic analysis of such spatial receptive fields obtained from an ensemble of AI neurons produced estimates of sound-source direction with an accuracy approaching that of a human listener (Jenison 1998). We also demonstrated that the gradient structure of the spatial receptive field that apparently underlies the ability of AI neuronal assemblies to encode sound-source direction is highly robust in the sense that it remains intact when continuous background noise is introduced into the sound field or when the level of stimulus is changed (Brugge et al. 1998).

The present report describes the sensitivities of AI cortical neurons to directional probe signals that are preceded in time by a conditioning sound originating from the same or different sound-source direction. In these experiments, we again used an approach based on the ability to deliver over earphones signals that mimic sounds coming from any chosen direction in free space (Chen et al. 1995; Reale et al. 1996, 1998; Wu et al. 1997). This configuration may be considered the simulation of a directly propagated sound and its echo. We first describe the time course of recovery of the response to a probe sound when the conditioning sound arises from the same or different directions in virtual acoustic space. We then go on to show that spatial receptive fields derived from responses to the probe signal, while altered by a conditioner, retain directional information.

METHODS

Animal use and care are in compliance with the “Guide for the Care and Use of Laboratory Animals” Publication No. 86-23 (revised 1985) of the National Institutes of Health and with the Animal Welfare Act of 1966 and its amendments of 1970 and 1976. Fifteen adult cats, with no sign of external or middle ear infection, were premedicated with acepromazine (0.2 mg/kg im) and ketamine (20 mg/kg im). A catheter was inserted into the femoral vein for intravenous drug administration and fluid replacement. Atropine sulfate (0.1 mg/kg sc), dexamethasone sodium (0.2 mg/kg iv), and procaine penicillin (300 K units im) were also administered before the animal was deeply anesthetized either with pentobarbital sodium (11 cats) or with halothane (4 cats). Pentobarbital sodium was administered intravenously (40 mg/kg). Halothane (0.8–1.8%) was administered with a carrier-gas mixture of oxygen (33%) and nitrous oxide (66%) through an endotracheal tube using a scavenged Verni-Trol vaporizer system and an anesthesia ventilator. Samples of inspiratory and expiratory air were drawn continuously from within the endotracheal tube and a respiratory gas analyzer (Ohmeda 5250) used to measure pulse rate, oxygen saturation, airway pressure, and concentrations of O₂, CO₂, N₂O, halothane, on a breath-by-breath basis. When halothane was employed, a muscle relaxant (pancuronium bromide, 0.15 mg/kg iv) was administered just before recordings began, if spontaneous respiration was irregular or otherwise compromised. Paralysis could be maintained throughout the experiment by supplemental doses of pancuronium. Muscle relaxation under halothane anesthesia, combined with careful monitoring of inspired and expired gases and vital signs, provided a highly stable long-term recording environment.

When the animal reached a surgical plane of anesthesia, the pinnae and other soft tissue were removed from the head. Hollow earpieces were inserted into the truncated ear canals, sealed in place, and connected to specially designed earphones. The transfer characteristics of the left- and right-ear sound-delivery systems were measured in vivo near the tympanic membrane. A chamber was cemented to the skull over the exposed left auditory cortex, filled with warm silicone oil, and sealed hydraulically with a glass plate on which a Davies-type microdrive was mounted. Action potentials were recorded extracellularly with tungsten-in-glass microelectrodes from single neurons in cortical area AI; their times-of-occurrence were measured with a 1-μs resolution and stored for off-line analyses.

Acoustic stimulus generation

Sound produced by a free-field source and recorded near the tympanic membrane of the cat is transformed in a direction-dependent manner by the pinna, head, and upper body structures (Musicant et al. 1990; Rice et al. 1992). The simulation of these transformed sounds as a function of sound-source direction constitutes a virtual acoustic space (VAS). A veridical model of VAS (Chen et al. 1995; Wu et al. 1997) was used to synthesize, in quasi-real time, transient signals for sound-source directions positioned in a spherical coordinate system (−180° to +180° azimuth, −36° to +90° elevation) and centered on the cat’s interaural axis. The VAS used was derived from a single cat. The intensity of any VAS signal was expressed simply as dB attenuation (dBA) relative to the maximum peak-to-peak amplitude for a particular sound source for that cat. Directional stimuli were impulsive transients (6.4-ms duration) that mimicked in their spectrum and time waveform sounds arriving from a source in free space. This simulated source generated either a broadband (3-dB corner frequencies at 800 Hz and 40 kHz) or narrowband (3-dB bandwidth = 2kHz) impulsive transient waveform that was realized as the impulse response of a linear-phase finite-impulse response filter. Narrowband sounds had their center frequency set equal to the characteristic frequency (CF) of the neuron and were employed for those neurons where broadband transient sounds did not evoke consistent responses. Digitally synthesized signals were compensated for the transmission characteristics of the sound-delivery system before D/A conversion. Tone burst stimuli delivered monaurally or binaurally were used to estimate the CF of a neuron and some response area features related to binaural interactions as described previously (Brugge et al. 1996). The partial tionotopic map obtained by repeated electrode penetrations made during the course of an experiment confirmed that the recordings were obtained from neurons in AI.

Measurements of directional sensitivity

Directional sensitivity was assessed using two stimulus paradigms. In one only a single directional probe sound was presented at a repetition period of 1–2 s to minimize the influence of a prior stimulus. In the other paradigm, a conditioning directional sound preceded the probe by a time interval ranging from a few milliseconds to hundreds of milliseconds. For each neuron isolated, we first obtained a virtual space receptive
field (VSRF or, simply, spatial receptive field) by delivering, in random order, a single probe sound from each of hundreds of VAS directions separated by 4.5° steps in azimuth and elevation, as described previously (Brugge et al. 1994, 1996; Reale et al. 1996, 1998). Each VSRF was constructed from the onset latency of the response to a sound at each of the VAS directions. Typically, we obtained VSRFs at several different intensities with one intensity being 20–30 dB above the threshold determined at the most sensitive region of the spatial receptive field. Second, we employed three variations of the conditioner-probe paradigm. In all three variations, the conditioner-probe pair was composed of two directional transient sounds, referred to as a lead and a lag, that were separated by an inter-sound interval (ISI) ranging from 0 to 400 ms. The repetition period between the lagging sound of one pair and the leading sound on the next pair was held constant at 1 s. In the first variation, responses were recorded to 40 repetitions of the two-sound stimulus that had both their leading and lagging directions coincident with the cell’s acoustic axis. Technically, the acoustic axis is that direction, for a single frequency, at which maximum pressure gain is recorded near the tympanum. The cell’s acoustic axis was operationally defined by the frequency corresponding to the CF of the neuron under study. Unless the ears are perfectly symmetrical, there is a different acoustic axis for each ear (Musicalt et al. 1990). The VAS constructed for these experiments employed symmetrical models for the left and right ears, and we further defined the cell’s acoustic axis to be on the side that yielded the lowest threshold. In the second variation, responses were recorded to 40 repetitions of a two-sound stimulus that had the lagging direction coincident with the cell’s acoustic axis combined with a different direction for the leading sound. For both of these variations, we typically employed ISIs of 0, 2, 4, 6, 8, 10, 20, 30, 40, 50, 100, 125, 200, 300, and 400 ms to derive recovery functions for both response magnitude (i.e., number of action potentials) and response latency (i.e., mean first spike latency). In the third variation of the conditioner-probe paradigm, we obtained VSRFs by combining a lead sound coincident with the cell’s acoustic axis with a lag sound delivered, in a random order and at a fixed ISI (50–400 ms), from each of hundreds of VAS directions separated by 4.5° steps in azimuth and elevation. Because of time limitations, spatial receptive fields derived from responses to these conditioned probe signals were often restricted to the frontal hemisphere. VSRFs were also modeled using a spherical approximation technique based on the so-called von Mises basis function (Jenison et al. 1998). This approach allowed us then to study the dependence of maximum-likelihood estimation performance on the spatial response properties of a population of AI neurons under conditions in which a single sound was presented in VAS and when the same sound was conditioned by a previous sound at different ISIs. The spherical approximation and ideal-observer methods employed in this study have been described in detail elsewhere (Jenison 1998, 2000; Jenison et al. 1998).

Typically, the discharge pattern to a conditioner-probe stimulus at ISIs greater than 20 ms could be unequivocally decomposed into responses attributable to the lead and lag sounds. Thus the spike count and response latency attributable to each sound of the two-sound stimulus were measured in nonoverlapping time windows of 10–20 ms. This situation obtained because under the conditions of these experiments, any effective stimulus evoked from a single AI neuron one spike or a short burst of action potentials whose times of occurrence typically spanned a range of less than 20 ms. Furthermore the majority of cells exhibited little or no spontaneous activity. When, however, the ISI was less than 20 ms, decomposition was not reliable and, thus the measurement window was widened to accommodate the total response to both transients.

RESULTS

Results presented in this paper were obtained from 50 isolated neurons in 15 cats drawn from a larger series of experiments in which we studied the VSRFs of several hundred neurons under VAS conditions (Brugge et al. 1994, 1996, 1998). The neurons reported in this paper responded to either broadband or narrowband directional transients and remained in contact with the electrode long enough to be studied parametrically and quantitatively for the effects of successive sounds on their spatial receptive field properties. The CFs of these neurons ranged from 7.3 to 28.6 kHz. Five of the 50 neurons were obtained under halothane anesthesia. Their recovery times after a conditioning stimulus were indistinguishable from those obtained under pentobarbital sodium anesthesia.

Interactions when successive signals arrive from the same direction on the cell’s acoustic axis

Figs. 1–4 illustrate data from two neurons obtained under the conditioner-probe paradigm (variation 1) where both the leading and lagging sounds arrived at the two ears from the same direction on the cell’s acoustic axis. The data shown are also representative of cells in our sample that were capable of responding consistently to both the leading and the lagging signal. The dot rasters (Figs. 1 and 3) show the patterns of action potentials generated by 40 repetitions of the two-sound stimulus for ISIs ranging from 0 to 400 ms. The accompanying recovery functions (Figs. 2 and 4) plot the change in response strength (normalized spike count) and response latency attributable to the leading and lagging sounds for each of the ISIs tested.

Figure 2 shows that for ISIs of 200 ms or greater the response to the lagging sound (●) resembled in both strength (A) and latency (C) the response to the lead (○). However, gradual reduction in ISI from 100 to 20 ms resulted in a rapid reduction and, eventually, total elimination of the response component attributable to the lagging sound. Effects on the discharge to the leading sound were unremarkable. At ISIs that spanned the range from 0 to 20 ms, we did not attempt to decompose the two-sound stimulus into lead and lag responses (see METHODS) but rather designated a joint response (Figs. 1 and 2, B and D). Strength of the joint response remained constant at about 50% of that obtained at the longest ISIs (300 and 400 ms) studied since at the latter ISIs each component of the two-sound stimulus produced nearly equivalent contributions to the joint response. Response latency showed little systematic change over the ISI range from 0 to 20 ms. These observations are consistent with assignment of the joint response to the leading sound with little or no contribution by the lagging sound. About 70% of neurons studied with the two-sound stimulus that had both the leading and lagging directions coincident with the cell’s acoustic axis exhibited recovery functions of the kind illustrated above. As shown in all of the dot rasters, the precision in response latency changed little with changes in ISI despite obvious changes in the absolute values of first-spike-latency and spike count.

The remaining 30% of cells in the group were distinguished by the shapes of their recovery functions at ISIs shorter than 20 ms. Data presented in Figs. 3 and 4 are representative of this group of neurons. Recovery functions for ISIs between 30 and 400 ms (Fig. 4, A and C), for both discharge strength and latency, are similar to those illustrated in Fig. 2. Across the range of shorter ISIs (Fig. 4, B and D), however, the recovery
function for response strength was nonmonotonically related to ISI, reaching a maximum at an interval of 8 ms. The fact that the response latency remained relatively constant under these conditions suggests that the increase in spike count occurred mainly in the later part of the discharge, a point made clear in the accompanying dot rasters (see Fig. 3). There, at an ISI of 20 ms, the discharge pattern consisted of a burst of three action potentials time-locked to the onset of the lead sound and preceding the onset of the lagging sound (▲). Reducing the ISI to 8 ms resulted in the emergence of a fourth time-locked action potential in the joint response. At still shorter ISIs this facilitation did not occur, accounting for the nonmonotonic character of the recovery function. Although this cell exhibited an increase in firing in this range of delay, a few neurons exhibited equally strong and systematic decreases (e.g., Fig. 6F).

Interactions when successive signals arrive from different directions

The question arises as to whether the relationships demonstrated when leading and lagging signals arrive from the same direction would be maintained when the leading sound originates from a direction different from that of the lagging sound (see METHODS, variation 2). Under such conditions, individual sounds in the stimulus pair may, when presented alone, evoke quite different response strengths and latencies reflecting their respective positions in the VSRF (see Brugge et al. 1996). In all 18 cases reported here, the lagging sound was placed on or near the cell’s acoustic axis, while the leading sound was usually located at a complementary direction in the opposite hemisphere. Thus the strength of the response to the leading sound alone could be the same or less than that exhibited by the
lagging sound when presented alone depending on the spatial spread of the VSRF at the intensity used. In 6 of the 18 cells, the leading sound failed to evoke a consistent response when presented alone. In five neurons, the response was less than 50% of that evoked by the lagging sound when presented alone, and in seven others the response exceeded 50%. In all cases the lagging sound on the cell’s acoustic axis always produced a robust time-locked discharge when presented without the conditioner.

Long recovery functions observed when both signals arrived successively from a single direction (variation 1) were also obtained when the signals arrived from different directions under the conditioner-probe paradigm (variation 2). Because of the similarity to results presented in Figs. 1–4, only one example of this will be illustrated here, representing the situation that arose when the leading sound, presented alone, evoked no spikes (Fig. 5). The VSRF for this neuron occupied the ipsilateral hemisphere (−180 to 0° azimuth), and the leading sound was positioned in the contralateral hemisphere (+36° azimuth and +27° elevation). Whereas the neuron showed no evidence of a response to this leading sound, the presence of the conditioner at this direction resulted in complete suppression of the response to the probe sound on the cell’s acoustic axis for ISIs between 2 and 50 ms. A partial recovery of response to the two-sound stimulus occurred at the ISI of 0 ms. Thus for this cell, and for all others where the lead and lag arrived from different directions and thus had differing response strengths and latencies, recovery functions measured in either strength or latency were similar, though not identical, to those observed when both sounds originated from the same direction on the cell’s acoustic axis.

Similarities and differences in recovery under the two conditions described in the preceding text are illustrated in Fig. 6. Here for all neurons in our sample the strength and latency recovery functions obtained when the two transients originated at the same direction (A, C, E, and G) are compared with the situation in which the lagging sound was on the cell’s acoustic axis and leading sound in the opposite hemisphere (B, D, F, and H). Long recovery times were observed regardless of the direction of the conditioner, the recovery functions were graded over a range of 50–200 ms or longer, with few exceptions complete suppression of the response to the lagging probe signal inevitability occurred at an ISI of 50 ms, and when affects were noted at ISIs <10 ms, they were most frequently of a facilitative or summative nature. Differences in the dynamic range of suppression were noted, however, as seen in the differences in slopes of the recovery functions shown in Fig. 6, A and B. To evaluate these differences, we obtained the ISI at which each strength recovery function fell by 50% (half-maximal point) and compared the distribution of this variable under the two conditioner-probe variations (Fig. 7). The median half-maximal value was 94.7 ms when both sounds arose from the cell’s acoustic axis but was significantly less than this (1-sample sign test, α = 0.004), 72.5 ms, when the leading sound arose from some other direction in VAS. Although the distributions are broad and overlapping, the results indicate that the strength of the interaction was related to the relative directions of the leading and lagging sounds within the VSRF.

Spatial receptive field dependence on a conditioning sound

To study the possible relationships between the relative directions of the leading and lagging sound on an AI directional response, we employed the third variation of the conditioner-probe paradigm: spatial receptive fields were obtained to probe sounds presented throughout VAS that were conditioned, at fixed ISIs, by a lead sound on the cell’s acoustic axis. This direction for the conditioner was chosen to ensure that the cell would respond robustly to the lead sound of every two-sound pair. Figure 8 illustrates results of this experiment on six AI neurons. VSRFs of these cells, plotted as orthographic projections, represent the major classes of VSRFs described previously in AI (Brugge et al. 1994, 1996). For each neuron isolated, we first obtained a VSRF by delivering, in a random order, the single probe sound from each of hundreds of VAS directions 20–30 dB above threshold (control), which was often sufficient for the field to cover a hemisphere. Each on the plot indicates the occurrence of a time-locked response to the stimulus at that direction. VSRFs shown below each control VSRF were obtained when the same probe sounds were each conditioned by a lead sound on the cell’s acoustic axis. The ISI (ms) employed is indicated to the top left of each VSRF.

Conditioned VSRFs were usually not demonstrably different from controls for ISIs of a few hundred milliseconds or longer. There was in all cases studied a decrease in the size of the conditioned VSRF when the ISI fell within the dynamic range of the neuron’s recovery function, as described earlier. The spatial receptive field was reduced in size at shorter ISIs where the decrease in the locations of effective directions tended to occur at the periphery of the VSRF. The ISI required to change the size of the VSRF varied from one neuron to the next, ranging from more than 200 ms (Fig. 8C) to about 70 ms (Fig. 8E). For most neurons studied, however, an interval of 75 or 85
ms was sufficient to severely reduce the size of the VSRF (e.g., Fig. 8, A, B, D, and F), and for all neurons in our sample at an ISI of 50 ms or less a response to the probe was hardly in evidence regardless of sound direction (e.g., Fig. 8, C and E). The size reduction of the VSRF occurred in such a way that the VSRF remained relatively circumscribed around the region in VAS that evoked the strongest response. This was commonly the region of the cell’s acoustic axis (e.g., Fig. 8A), although, as illustrated in Fig. 8B, this was not always the case. The ISI-dependent reduction in size of the VSRF illustrated here is consistent with our observations of the time course of recovery in response strength made earlier with successive signals arriving from just one or two directions (Figs. 1–7). This effect of a previous sound on the VSRF was observed for all of the receptive field classes described in our earlier work (Brugge et al. 1996).

Previous results indicated that AI neurons spatial receptive fields exhibit systematic gradients of response latency or response strength (Brugge et al. 1994, 1996) that can be quantitatively modeled using spherical basis functions (Jenison et al. 1998). In the frontal hemisphere, gradients of increasing latency or decreasing strength can radiate from a central location in the field, often near the cell’s acoustic axis. Our recorded sample of spatial receptive fields were obtained at a relatively high spatial resolution with each of hundreds of sample directions tested only one time. Thus we used response latency rather than spike rate as the measure of response strength as these two variables are typically highly correlated (Brugge et al. 1996; see also Phillips 1989). Moreover response latency itself can provide information about sound direction (Jenison 1998).

Figure 9 illustrates the distribution of response latency for each of the VSRFs illustrated in Fig. 8. For each stimulus condition, response latency was restricted to several millisecond intervals.
onds around a mean latency. The mean latency tended to lengthen with decreasing ISI. In one or two cases, there were two peaks in the distribution. In previous studies, we demonstrated that not only was there a functional gradient in the VSRSF but that a gradient was maintained in the face of changes in signal intensity or in the presence of continuous background noise (Brugge et al. 1998). The question here is whether systematic VSRSF gradients are also delineated in the presence of a conditioning sound that is positioned to fall near the most sensitive direction in the cell’s spatial receptive field.

VSRSFs were constructed to show the spatial distribution of response latency when a single probe was employed (control) and when this probe was preceded, at a fixed ISI, by a conditioning sound positioned on the cell’s acoustic axis. A set of these VSRSFs, the frontal-field portions of which were illustrated in Fig. 8A, are illustrated here as quartic-authalic equal area maps (Fig. 10). This map projection displays in two dimensions the entire auditory space surrounding the animal and minimizes distortion in the frontal hemisphere. Empirical data (left) from this representative neuron were used to construct a VSRSF based on first-spike-latency (color coded) for each ISI. Nonresponsive directions are colored black. The empirical data were also modeled with spherical basis functions (right) to provide a mathematical characterization of the systematic latency organization within the spatial receptive field (Jenison et al. 1998). Using this approach the VSRSF can be described by a continuous function, a feature we have taken advantage of in deriving azimuth functions from each of the modeled VSRSFs.

From Fig. 10 we see first that under control conditions, when the stimulus consisted of but a single transient signal, there was an orderly spatial representation of response latency. The directions in VAS that were associated with shortest latency tended to cluster, and this aggregation of shortest latency was surrounded by directions for which latency became progressively longer. This general pattern was observed among VSRSFs from different AI neurons, although the exact shapes of iso-latency contours may vary (see also Brugge et al. 1996).

Second, an orderly gradient structure remained even when the size of the VSRSF was demonstrably reduced by the presence of a conditioning sound falling on the most sensitive region of the VSRSF and within a few hundred milliseconds of the probe. As we already showed in the preceding text (see Fig. 9A), the absolute latency and the its distribution may shift with changes in ISI, and these shifts are reflected in comparisons of gradient patterns between the control VSRSF and any conditioned VSRSF. Nevertheless the center of the VSRSFs remained at or near that of the control with gradients in response latency radiating from it. Third, the model used was a reasonable representation of the empirical VSRSF.

**Gradients of response latency as observed in derived azimuth functions**

Models of the kind shown above in Fig. 10 can reveal in fine detail the spatial gradients of response latency that are exhibited by AI spatial receptive fields. From modeled VSRSFs, we were able to derive plots of response latency as a function of stimulus azimuth at a fixed elevation (azimuth functions). In previous studies (Brugge et al. 1996, 1998), we showed that such graphs exhibited steep functional gradients in the receptive field, which we interpreted as being a key response feature underlying directional acuity (Jenison 1998). In Fig. 11 we illustrate families of azimuth functions obtained from eight neurons. The fixed elevation chosen for each passed through the cell’s acoustic axis. The parameter is ISI. Dashed lines represent data derived from modeled control VSRSFs where the single-sound paradigm was employed. Solid curves represent data, at different ISIs, from modeled VSRSFs in the conditioner-probe paradigm. Figure 11, A–G, illustrates data from seven neurons plotted only for the frontal acoustic hemisphere; the curves in Fig. 11H are continued into the right rear acoustic quadrant. Figure 11, A and B, was derived from VSRSFs shown in Figs. 8 and 10. In all cases shown, as well as in those not illustrated here, response latency changed systematically with changes in stimulus direction along the azimuth. Response latency tended to be shortest for directions in the right acoustic hemisphere (contralateral to the cortex under study) and became progressively longer for directions on either side of this. We note especially that the greatest changes in latency tended to occur with changes in azimuth direction across the midline, between about +18° and −18°.

**Maximum likelihood estimation of direction from ensemble responses**

It was shown previously that a theoretical ideal observer can be derived to estimate sound-source direction from an ensemble of AI neurons whose empirically measured VSRSFs have been described analytically (Jenison 1998, 2000). In this approach, the systematic fine structure of first-spike latency in a VSRSF was characterized by approximation with spherical basis functions, and the unsystematic noise was isolated in the residuals. Examination of the residuals supported a linear model for the dependence of variance on first-spike latency. Each neuron in an ensemble of AI neurons was then reasoned to possess a Gaussian probability density function that utilized both the spherical basis function model of the spatial receptive field and the linear model of variance (Jenison et al. 1998). Thus a single presentation of a sound from a source at a particular direction elicits from this AI ensemble a correspond-
ing population of response latency values that is conditioned on only two parameters: azimuth and elevation. The ideal observer estimates the azimuth-elevation pair that maximizes the likelihood function, which under the assumption of independent noise between the cells, is the joint product of the probability density functions of all the neurons in the ensemble. Importantly, a second presentation of the same sound at the same source direction would typically result in a different set of response latency values and, hence, a slightly different estimate of the parameter azimuth-elevation pair by the ideal observer. A lower bound on the variance of these repeated estimates was provided by calculating the Cramer-Rao lower bound (CRLB). The CRLB was then calculated for different sound-source directions and different ensemble sizes.

A similar simulation was employed here. A set of VSRFs from one single neuron were each modeled (e.g., see Fig. 10, right) using spherical basis functions (Jenison et al. 1998). The set consisted of a control VSRF and conditioned VSRFs (one for each ISI). The VSRFs of this modeled set were then replicated using the empirical distribution of receptive field centroids (Fig. 12A) mapped previously by Brugge et al. (1996) thereby producing a 65-cell neural ensemble. The simulation is not expected to be critically dependent on this number of neurons, since near-asymptotic CRLBs have been estimated with far fewer neurons (Jenison 1998). The results of this simulation are shown in Fig. 12B. There, ○ represents results from the control VSRFs and ■, ♦, ▼, ▲, and ● are from conditioned VSRFs, at the designated ISI. The ordinate shows the minimal error (square root of the CRLB) calculated with the ideal observer analysis; the abscissa marks those directions along the azimuth (for a fixed elevation) at which the simulation was executed. For ISIs of a few hundred milliseconds, error functions were within about two degrees of the control function. At shorter ISIs the outcome of the simulation showed that there was no major degradation (re control) in the ensemble performance. On the contrary, performance of the conditioned ensemble was often slightly better than the control at the shortest ISIs modeled. This outcome is a direct consequence of

FIG. 5. See legend of Fig. 1. In this case, however, the leading and lagging sounds were at 2 complementary directions in VAS.
the maintenance of internal structure in the contracted spatial receptive fields and of the assumption of independence among these neurons. In summary, these simulation results suggest that the directional precision provided by ensembles of AI neurons would be largely immune to competing transient sounds within the range of ISIs investigated.

Interactions exhibited while listening with one ear

For any sound-source direction in VAS, the signals delivered to the right and left ears are appropriate in interaural time, intensity, and spectrum for that particular direction in the free field. All recovery functions discussed to this point were studied under these binaural conditions. However, there is evidence at the level of the midbrain to indicate that the neural interactions observed when sounds are delivered binaurally in rapid succession may also be observed with monaural delivery and therefore lacking these binaural cues (Yin 1994). Although this question was not central to the current study, in four cases, we obtained recovery functions under both binaural and monaural listening conditions across the full set of ISI values. In these four cases, monaural listening to the directional signal was achieved simply by delivering sound only to the ear contralateral to the cortex under study. Using this monaural adaptation, two of the neurons were studied with variation 1 (i.e., both lead and lag directions on the cell’s acoustic axis) and the other two cells with variation 2 (i.e., only lagging direction on the cell’s acoustic axis). Although the number of such comparisons between monaural and binaural stimulus conditions are few, the results seemed unequivocal in that under both conditions the shapes of recovery functions were similar to those obtained under binaural listening conditions.

DISCUSSION

In the present experiments, we employed transient conditioner and probe sounds that contained temporal and spectral properties appropriate for their respective sound-source directions in the free field. We found that, without exception, the response of an AI neuron to the probe sound was suppressed if preceded in time by a conditioning sound from the same sound source located at the same or different direction within the cell’s spatial receptive field. Suppression was a graded function of ISI; it was weakest for ISIs of a few hundred milliseconds and became progressively stronger as the ISI was decreased systematically. For all cells, complete suppression was observed when the interval between conditioner and probe was reduced to about 50 ms. The occurrence of suppression need
not depend on the leading sound evoking spike discharges from the cell. However, the dynamic range for suppression was, on the average, shorter when the two sounds arose from two widely spaced directions as compared with the same direction. Long recovery times obtained under monaural listening conditions were within the range of those observed under binaural conditions. For some neurons, there was a response to a conditioner-probe stimulus at ISIs between about 4 and 10 ms that was not equivalent to the response to the conditioner alone.

Influence of a conditioning sound on recovery time

Recovery functions we recorded were similar in many ways to those obtained by others from field AI of the anesthetized cat based on evoked-potential responses to nondirectional click pairs delivered monaurally or binaurally (Borsanyi 1964; Hocherman and Gilat 1981; Perl and Casby 1954; Rosenblith 1950; Rosenblith and Rosenzweig 1951; Rosenzweig 1951, 1954; Rosenzweig and Rosenblith 1950, 1953). We did not observe, however, the cyclic recovery functions reported by Rosenzweig and Rosenblith (1953). The recovery times we obtained also agree with those reported in forward-masking studies of single AI neurons using nondirectional monaural or binaural clicks (Serkov and Yanovskii 1970), or tone bursts (Brosch and Schreiner 1997; Calford and Semple 1995; Schreiner et al. 1997). They are also consistent with the time course of suppression of spontaneous activity following the transient response to a free-field click (Eggermont 1991).

The mechanisms that underlie the relatively long recovery to an acoustic transient are not limited by the refractory period of the neuron, for AI cells are quite capable of firing bursts of spikes with interspike intervals of 1–2 ms, and some can entrain to repetitive stimulation at repetition rates approaching 800/s (de Ribaupierre and Goldstein 1972). Recovery time of an AI neuron is also not limited by the precision in spike timing. Phillips’ (1989) studies of AI neurons in cats under pentobarbital sodium anesthesia showed that for brief tone pulses repeated at rates where spike rate and spike entrainment declined toward zero, the precision of spike timing was easily sufficient to sustain perfect entrainment for frequencies of 60–100 Hz. Our results also indicate that whereas first spike latency typically lengthens during the suppression period, the timing of the spikes remains quite precise. Phillips (1989) suggested that other factors, such as adaptation, must contribute to the high-frequency cutoff of the response of an AI neuron to repeated acoustic transients. Eggermont’s (1999) more formal analysis of the low-pass characteristics of AI neurons includes adaptation in a model based on presynaptic facilitation and depression. Calford and Semple (1995) hypothesized that forward masking is the result of inhibitory mechanisms operating at the level of the cortex or at lower levels in the auditory pathway. Cortical inhibition is further suggested by the earlier results of Serkov and Yanovskii (1971, 1972) from intracellular recordings from AI neurons in the unanesthetized and curarized cat. Here it was shown that an acoustic click or a shock to the auditory thalamic radiations resulted in a long-lasting hyperpolarization. The presence of a long-lasting hyperpolarization was also shown not to require a preceding excitatory event, a finding that is in full agreement with our results. For the great majority of cells recorded under these conditions, the duration of hyperpolarization was found to be 30–200 ms, with the most common duration being in the range of 80–100 ms. This time interval overlaps the dynamic range of suppression we observed using directional transients.

The recovery time of an AI neuron following the presentation of a transient sound in space depended on the relative directions from which the conditioning and probe stimuli arose. On average, the longest suppression was caused by conditioning stimuli arising from the most sensitive region of the spatial receptive field, which is typically the cell’s acoustic axis. Here stimulus amplitude is greatest at the CF of the neuron under study. Hence, this area is often a spatial focus of highest response magnitude and shortest response latency (Brugge et al. 1996). These results complement those of Brosch and Schreiner (1997), which showed that the recovery time of an AI neuron following a tone burst depended on the relative positions of the conditioning and probe stimuli in the neuron’s frequency-intensity response area; the longest suppression occurred when the conditioner (masker) was at the neuron’s CF. Our results are also in accord with the earlier findings that recovery of responsiveness of AI to a probe click, as measured by changes in membrane potential (Serkov and Yanovskii 1971, 1972) and the magnitude of the evoked potential (Rosenzweig 1954; Rosenzweig and Rosenblith 1953), depends on the intensity of a conditioning click and hence on the amplitude of the response evoked by that signal. These previous findings in cortical area AI of a relatively long recovery period after the presentation of nondirectional clicks, tones, and noise or electrical stimulation of geniculo-cortical radiations, coupled with our new data showing very similar recovery times in response to directional transients, lead us to conclude that the same neural circuits in AI may be engaged in forward masking, in processing temporal sequences, and in integrating directional sounds in reverberant spaces.

The question naturally arises concerning the possible effects of anesthesia on the AI recovery process and hence on the
recovery time. In cats anesthetized with sodium pentobarbital (Borsanyi 1964; Etholm et al. 1976) or with chloralose (Borsanyi 1964), the recovery time of the averaged click-evoked potential recorded in AI was found to be lengthened by tens of milliseconds as compared with the awake animal. Aitkin and Dunlop (1968, 1969) and Etholm et al. (1976) found a similar anesthesia effect on evoked potentials recorded in the medial geniculate body of the cat. They noted a comparatively smaller effect in the inferior colliculus (IC); a finding confirmed in single unit studies in the rabbit IC by Fitzpatrick et al. (1995). Ketamine and pentobarbital sodium were shown to result in equivalent estimates of recovery times (Brosch and Schreiner 1997; Calford and Semple 1995). Although AI recovery times may be lengthened by general anesthesia, recovery times have also been reported to range as high as 300–700 ms in the unanesthetized cat (Serkov and Yanovskii 1970) and rabbit (Fitzpatrick et al. 1997, 1999). In the auditory cortex of the awake human (Howard et al. 2000; Lee et al. 1984; Liegeois-Chauvel 1991; Puletti and Celesia 1970) and the monkey (Lu et al. 1999), the recovery functions derived from click-evoked responses are very similar to those we obtained from single neurons in AI of the anesthetized cat. Furthermore the results of intracellular recordings of AI neurons in the unanesthetized (but paralyzed) cat have shown that acoustic transients commonly evoke inhibition lasting tens to hundreds of milliseconds (Serkov and Yanovskii 1971, 1972). Regardless of the anesthetic state of the animal, AI recovery times are uniformly longer than those recorded in subcortical auditory regions.

![Fig. 8](http://jn.physiology.org/)

**FIG. 8.** Virtual space receptive fields (VSRFs; frontal acoustic hemifield only) from 6 neurons (A–F) obtained from responses to a single transient stimulus presented once at each direction (control) or to the lagging sound of a pair of transients presented once at each direction at different ISIs (shown in milliseconds to the top left of each VSRF). Simple orthographic projection. Black box, time-locked response of the neuron to the lead sound (control) or to the lag sound at the designated ISI at that direction. Data collected at a spatial resolution of 4.5° (C and D) or 9° (A, B, E, and F).
including the MGB (Aitkin and Dunlop 1968, 1969; Aitkin et al. 1966; Etholm et al. 1976), superior olivary complex (Fitzpatrick et al. 1995, 1999), IC (Fitzpatrick et al. 1995, 1999; Litovsky and Yin 1998a,b; Yin 1994), cochlear nuclei (Fitzpatrick et al. 1999; Kaltenbach et al. 1993; Parham et al. 1998; Wickesberg 1996), and auditory nerve (Fitzpatrick et al. 1999; Parham et al. 1996). Creutzfeldt et al. (1980) demonstrated directly in the unanesthetized guinea pig that auditory cortical neurons ceased to follow frequencies above 20 Hz, whereas thalamocortical neurons projecting on them could follow at frequencies up to five time higher. Thus cortical limits in the temporal processing of repetitive transients appear to result from the progressively longer time constants of inhibitory mechanisms that become evident at successive stages in the ascending auditory pathway.

**Minimal recovery time**

Despite the considerable amount of electrophysiological evidence that characterizes auditory cortex as having long time constants, there are also data indicating that a certain proportion of AI cells may have recovery time constants that are relatively short. Serkov and Yanovskii (1970) reported that in the unanesthetized cat recovery times of AI evoked potentials could be as short as 3 ms, although the greatest majority of cells they recorded exhibited recovery times 80 ms, with some as long as 700 ms. Fitzpatrick et al. (1997, 1999) also showed that auditory cortical recovery times in the unanesthetized rabbit were typically long, although some were as short as 2–3 ms. More recently, Mickey and his colleagues (1999, 2000) reported that neurons in fields AI and AII of the cat cortex exhibit interactions in the response to clicks arriving from two directions with ISIs of less than 1 ms. We did not study ISIs within a 1-ms time frame. In certain AI cells, however, we did observe increases or decreases in firing strength associated with changes in discharge pattern at ISIs between about 4 and 10 ms. A similar finding has been reported in auditory cortex of the awake monkey (Lu et al. 1999). Schreiner and his colleagues have shown that AI cortex is

![Graph showing distribution of 1st-spike latency at different ISIs across the VSRFs of 6 neurons.](http://jn.physiology.org/)

**FIG. 9.** Distribution of 1st-spike latency at different ISIs across the VSRFs of the 6 neurons illustrated in Fig. 8. ISI in milliseconds, number of spikes (N), mean 1st spike latency (M), and standard deviation (SD) given on each panel.
functionally segregated along a number of acoustic dimensions (for review, see Schreiner 1998), and thus it may be that neurons with short time constants are confined to areas not sampled in our experiments. It is also possible that such neurons occupy layers of cortex that are usually relatively silent under the general anesthesia used here.

Thus there appears to be a population of AI neurons sensitive to short temporal intervals as well as to intervals in the range of tens to hundreds of milliseconds. It may be that these populations of cells are engaged in processes associated with the precedence phenomena of fusion and lag discrimination suppression (Litovsky et al. 1999) and summing localization (Mickey et al. 1999). A possible role of auditory cortex in precedence phenomena is also suggested by results from lesion-behavior studies (Cranford 1982; Cranford et al. 1971; Masterton and Diamond 1964; Whitfield et al. 1972, 1978). Several studies have shown that cats discriminate interaural time differences of a few milliseconds and thus may experience the precedence effect (Cranford 1982; Cranford and Oberholtzer 1976) and that lesions of auditory cortex affect localization performance that depends on processing short time intervals (Cranford et al. 1971; Masterton and Diamond 1964; Whitfield et al. 1972, 1978). Recently, Mickey and Middlebrooks (2000) reported on auditory cortical neurons whose response properties may help account for such behavior.
Influence of a conditioning sound on spatial receptive field

In the present experiments, when a conditioning sound was positioned near the most sensitive directions in the neuron’s spatial receptive field (typically but not always the cell’s acoustic axis), successive probe sounds were more likely to be suppressed when their sound-source directions were farther from, rather than nearer to, the direction of the conditioner. Thus the spatial receptive field, when conditioned at ISIs from 50 to 400 ms, usually appears to contract about the direction of the conditioner.

In human psychophysical studies, listeners can easily resolve the presence of two successive sounds arising within a few tens to hundreds of milliseconds of each other regardless of their directions in space. Nevertheless, the leading sound can have a powerful residual influence on the spatial acuity (Grantham 1986; Perrott and Pacheco 1989; Tollin and Henning 1998) or perception (Hari 1995; Stellmack et al. 1997; Thurlow et al. 1965) of the lagging sound in this time interval. These lingering perceptual effects might be related to the significant changes in absolute response latency and firing probability that we observed routinely in the spatial receptive fields of AI cells following a conditioning signal. Previously we reported that the VSRFs of a considerable proportion of recorded AI neurons are characterized by spatial gradients of response latency and/or firing probability (Brugge et al. 1994, 1996). An ideal observer model based on an ensemble of AI neurons predicted that sufficient information was provided by these functional gradients to account for auditory spatial acuity of both of cat and human in anechoic space (Jenison 1998, 2000; Jenison et al. 1997). In the current study, both firing probability and response latency were altered by the conditioning stimulus, although spatial gradients based on these response measures were still recognized in the spatial receptive field. An ideal observer analysis of a simulated ensemble of these AI neurons showed that high spatial acuity was preserved for conditioning ISIs ranging from tens to hundreds of milliseconds. Functional gradients are also maintained when the size and shape of the VSRF field was altered by changes in stimulus intensity (Brugge et al. 1996, 1998) or by the presence continuous background noise (Brugge et al. 1998). In other words, the gradient structure of an AI spatial receptive field is highly robust. It appears, therefore that the VSRF gradients maintained by AI neurons in the face of these conditioning or competing sounds may be sufficient to support much of a listener’s ability to perceive a sound in space and judge its direction.

Although lesions of auditory cortex disrupt sound localization performance based on detection of small interaural time differences, no study of this kind in experimental animals has taken up the question of possible effects on those aspects of spatial hearing that involve long interstimulus time intervals. Lesions of auditory cortex affect localization performance in a
way that suggests that they actually disrupted the animal’s organization of auditory space (Heffner and Heffner 1990; Wegener 1964). Location in space appears to be no longer a part of this animal’s perception of the external environment (Masterton and Diamond 1964), and therefore to such an animal, a sound source does not have a right or wrong direction, it simply has no direction (Whittle 1977). Thus it may be that primary auditory cortex is part of a forebrain circuit that plays a role, not only in localizing the source of a sound per se but in the animal’s perception of its external acoustic environment.

This work was supported by National Institutes of Health Grants DC-00116 and HD-03352.

REFERENCES


ROSENZWEIG MR AND ROSENBLITH WA. Responses to auditory stimuli at the cochlea and at the auditory cortex. Psychol. Monogr 67: 1–26, 1953.


