Azimuth Coding in Primary Auditory Cortex of the Cat. I. Spike
Synchrony Versus Spike Count Representations

Jos J. Eggermont and Jennifer E. Mossop
Departments of Physiology and Biophysics, and Psychology, The University of Calgary, Calgary, Alberta T2N 1N4, Canada

Eggermont, Jos J. and Jennifer E. Mossop. Azimuth coding in primary auditory cortex of the cat. I. Spike synchrony versus spike count representations. J. Neurophysiol. 80: 2133–2150, 1998. The neural representation of sound azimuth in auditory cortex most often is considered to be average firing rate, and azimuth tuning curves based thereupon appear to be rather broad. Coincident firings of simultaneously recorded neurons could provide an improved representation of sound azimuth compared with that contained in the firing rate in either of the units. In the present study, a comparison was made between local field potentials and several measures based on unit firing rate and coincident firing with respect to their azimuth-tuning curve bandwidth. Noise bursts, covering a 60-dB intensity range, were presented from nine speakers arranged in a semicircular array with a radius of 55 cm in the animal’s frontal half field. At threshold intensities, all local field potential (LFP) recordings showed preferences for contralateral azimuths. Multunit recordings showed in 74% a threshold for contralateral azimuths, in 16% for frontal azimuths, and in only 5% showed an ipsilateral threshold. The remaining 5% were not spatially tuned. Representations for directionally sensitive units based on coincident firings provided significantly sharper tuning (50–60° bandwidth at 25 dB above the lowest threshold) than those based on firing rate (bandwidths of 80–90°). The ability to predict sound azimuth from the directional information contained in the neural population activity was simulated by combining the responses of the 102 single units. Peak firing rates and coincident firings with LFPs at the preferred azimuth for each unit were used to construct a population vector. At stimulus levels of ≥40 dB SPL, the prediction function was sigmoidal with the predicted frontal azimuth coinciding with the frontal speaker position. Sound azimuths >45° from the midline all resulted in predicted values of ~90 or 90°, respectively. No differences were observed in the performance of the prediction based on firing rate or coincident firings for these intensities. This suggests that although coincident firings produce narrower azimuth tuning curves, the information contained in the overall neural population does not increase compared with that contained in a firing rate representation. The relatively poor performance of the population vector further suggests that primary auditory cortex does not code sound azimuth by a globally distributed measure of peak firing rate or coincident firing.

INTRODUCTION

Cats can localize sounds in the horizontal plane presented within 20° from the midline with an error of at most 10° regardless of whether transient stimuli or long-duration sounds are used. Behavioral accuracy is greatest for sounds in the midline where the error is as low as 2° (Populin and Yin 1998). Unilateral lesions of primary auditory cortex (AI) in cat (Jenkins and Merzenich 1984) and ferret (Kavanagh and Kelly 1987), result in either exclusively or predominantly contralateral deficits under identical test conditions. Lesions that spare AI or a small iso-frequency patch in AI allow sound localization of brief sounds, suggesting that AI is both sufficient and necessary for unimpaired sound localization (Jenkins and Merzenich 1984). Unilateral lesions preserve normal head and eye orientation to the location of a sound source (Beitel and Kaas 1993). This suggests that, at least in cat and ferret, the primary auditory cortex is needed to localize brief sounds presented in the contralateral hemifield or participate in more complex localization behavior.

Single-unit firing rate representation

From several detailed single (SU) and multiunit (MU) firing rate studies in cat primary auditory cortex (Clarey et al. 1994; Imig et al. 1990; Middlebrooks and Pettigrew 1981; Rajan et al. 1990a), it is now well established that there is no map of azimuth in AI that is based on firing rate. Azimuth representation in AI, nevertheless, appears less stereotyped than in the inferior colliculus (IC) (Aitkin and Jones 1982) and in the medial geniculate body (MGB) (Barone et al. 1996; Clarey et al. 1995). Responses to sound in AI are represented by SUs with omnidirectional, contralateral, ipsilateral, or multipeaked response preferences that are strongly intensity dependent. For example, receptive fields that are omnidirectional at high-intensity levels become generally tuned to sound in a restricted region of the contralateral hemifield at intensities within 10 dB of threshold at best azimuth (Brugge et al. 1996; Middlebrooks and Pettigrew 1981; Rajan et al. 1990a). This level dependence may result from the directionality of the pinna, which boosts pure tones, in the 10- to 30-kHz range, ≥30 dB when their azimuth is aligned properly with the pinna axis (Phillips et al. 1982). Another contributing factor may be the presence of mid-frequency (8–16 kHz) notches in the sound spectrum at the eardrum (Irvine 1987; May and Huang 1996; Rice et al. 1992) that appear to shift to lower frequency for azimuth shifts from ipsilateral to midline but stay relatively constant at contralateral sound locations. Because spatial response areas for wideband noise bursts and tones at the characteristic frequency (CF) are generally very similar (Rajan et al. 1990a), the directional sensitivity appears to be determined largely by pinna effects and/or by the noise components around CF.

The four studies that provided the vast majority of data on azimuth sensitivity in AI differed in the proportion of omnidirectional units that were found. This proportion
ranged from about half (Middlebrooks and Pettigrew 1981) to ~20% (Brugge et al. 1996; Imig et al. 1990; Rajan et al. 1990a). Anesthesia differed in these studies: ketamine only (Middlebrooks and Pettigrew 1981), pentobarbital initially followed and maintained by ketamine (Rajan et al. 1990a) or pentobarbital only (Brugge et al. 1996; Imig et al. 1990). The type of sound used also differed: tonal stimuli (Middlebrooks and Pettigrew 1981), noise bursts (Imig et al. 1990), or both (Rajan et al. 1990a). In the latter study, no major differences in azimuth tuning preference were found for noise and tonal stimuli. It is thus possible that a difference in anesthesia (specifically the presence or absence of pentobarbital) could account for some of the differences.

Directional sensitivity under open field stimulation conditions for units with CFs >3–4 kHz is likely due to interaural intensity (IID) differences and pinna effects. Explicit studies to test the IID contribution by plugging one ear and comparing the cortical unit's receptive fields to the unplugged condition (Samson et al. 1993, 1994), showed that about one-quarter of the azimuth sensitive cells had monaural directional fields to noise-burst stimulation, likely resulting from spectral cues. Directional receptive fields obtained for noise bursts and CF-tone bursts had generally the same azimuth preference but showed a tendency to be more selective for noise (Clarey et al. 1995; Rajan et al. 1990a).

Units with different azimuth-sensitive response properties appeared to be spatially segregated (Middlebrooks and Pettigrew 1981; Rajan et al. 1990b), and sometimes repeated clusters of a particular response type, but not the omnidirectional one, could be found along isofrequency strips (Rajan et al. 1990b). About three-quarters of the radial penetrations showed a homogeneity for azimuth preference but slightly less than half showed this homogeneity also for the level dependence. MU and SU preferred azimuth was generally the same, albeit that they often showed different suprathreshold response properties (Clarey et al. 1994). This suggests a distributed representation of sound azimuth in AI and the need to postulate a population code.

Because spatial receptive fields in primary auditory cortex are mostly broad and level dependent (Brugge et al. 1996; Imig et al. 1990; Middlebrooks and Pettigrew 1981; Rajan et al. 1990a) and because AI is indispensable for sound localization in cats (Jenkins and Merzenich 1984), the role of AI for sound localization in cats could more likely be a cognitive one (Beitel and Kaas 1993) rather than the more reflexive role of the superior colliculus (SC). Aitkin and Jones (1992) also suggested that the processing of auditory azimuth information from the IC by the SC (where maps of azimuth are found) and by AI is largely a sign of parallel processing. A much larger proportion of azimuth-sensitive cells was found in AI than in IC, suggesting an increased importance of higher auditory regions for sound localization in cats (Aitkin and Jones 1992; Barone et al. 1996). Because the AI projects to the SC (Winer 1992), a corticofugal modulation of neural activity in the IC or SC by the AI cannot be ruled out (Zhang et al. 1997).

Neural firing synchrony representation

The potential role for synchronous activity between adjacent neurons in encoding sound location in auditory cortex was first suggested by Bloom and Gerstein (1984) and tested by Ahissar et al. (1992). Synchrony of firings in auditory cortex occurs largely because the neurons respond in time-locked fashion to certain aspects, e.g., the onset or AM, of the stimulus. Synchrony in this case refers to near simultaneous firings of two neurons, say within 10–15 ms of each other so that the cross-correlation peak is fairly narrow. Neural synchrony (Eggermont 1994) is broader defined than coincident firings, which usually is considered to require firings within 1 ms. Such coincident firings have been proposed for brain stem and midbrain sound localization mechanisms that convert interaural time differences (compensated by an array of different neural path lengths to allow coincidence detection) to spatially distinct activation (Yin and Kuwada 1984). In some of Bloom and Gerstein's (1984) recordings, only those that were indicative of direct neural interaction, the shape and strength of the correlograms appeared to be more sensitive to sound location than the SU firing rate. The proportion of those direct interactions was, however, very low. Ahissar et al. (1992) found in awake macaque monkeys that, for stimulation at 70–80 dB SPL, ~62% of neurons showed a modulation of firing rate as a function of sound azimuth; 60% of these neurons preferred contralateral stimuli. After correction for stimulus-induced synchrony with the shift-predictor, the shapes of a few correlograms were dependent on sound azimuth. This shift predictor (Perkel et al. 1967) is the normalized product of the poststimulus time histograms (PSTHs) of the two neurons involved and as such does not preserve the firing time relationship of the units but only their average joint firing timing properties with respect to the stimulus. However, one can imagine that correcting for the stimulus-induced synchrony removes the very aspects of neural synchrony that the animal needs to make an orientation or localization response. Furthermore because the time of stimulus presentation is unknown to the animal, it will be unable to calculate a shift predictor and thus to separate that part of the interneural synchrony that is not the result of firings that are synchronous with stimulus onset. A more elaborate survey of the potential of neural synchrony, without correction for stimulus induced synchrony, to represent sound azimuth therefore is needed.

The firing rate and firing synchrony representations studied in primary auditory cortex so far are local: they comprise the pooled firing rate recorded on one electrode or the interneural synchrony on a single electrode or across closely spaced dual electrodes. However, coding of sound azimuth in AI might be based on the activity of an ensemble of neurons distributed across AI, as demonstrated for the auditory midbrain (Knudsen et al. 1987). A recent study in the anterior ectosylvian sulcus (AES) (Middlebrooks et al. 1994) has suggested that a temporal code could be more appropriate than a rate code for revealing spatial tuning. They noted that a dominant feature of spike timing that varied with location was the overall latency. Population coding based on temporal information in the spike trains and extracted by a one-layer linear neural network appears to provide a crude but monotonic representation of azimuth in the AES (Middlebrooks et al. 1994). More recently, Middlebrooks and Xu (1996) showed that firing rate was still the most important factor in determining the directional in-
formulation in panoramic responses and that temporal patterning was the next most important factor.

Several other methods, distinguished by the way neural activity is combined, may reveal a population code of sound azimuth. The most commonly employed model is that of combining the firing rates of individual neurons (Abbott 1994). Georgopoulos and colleagues (1986, 1988) have pioneered the use of such population rate codes in primate motor cortex. Neuronal activity was recorded in primate motor cortex simultaneously with the direction of arm movements in three-dimensional space. The discharge rate of 84% of the cells recorded varied in an orderly fashion with the direction of movement. The discharge rate of a single cell was highest for movements in a certain direction and decreased progressively with movements in other directions roughly as a function of the cosine of the angle formed by the direction of movement and the cell’s preferred direction. Individual units thus were generally broadly tuned for direction: the cosine tuning curve has a bandwidth at half-peak amplitude of 90°. This population-vector coding model assumed that for a particular movement direction, each cell makes a contribution in the cell’s preferred direction with a weight proportional to the change in the cell’s discharge rate associated with the particular direction of movement. The vector sum of all these contributions is the neural population vector, the readout of the population code, which points in the direction of movement in space.

Modeling the way a particular neural structure might potentially perform its task in no way proves that this procedure actually is used in the nervous system. However, it is a way to estimate the amount of information about the location of a sound source that is present in the activity of an ensemble of neurons. The estimate of sound azimuth produced by vector addition of firing rates can be computed by the integrative action of individual neurons or groups of neurons. The population-vector method can be applied to the average firing rate in a specified time window, to changes in firing rate from spontaneous or average activity, to normalized firing rates, etc. Therefore the magnitude of the constructed direction vector is strongly dependent on the assumptions made. In contrast, the direction of the population vector is unaffected by such linear transformations (Salinas and Abbott 1994).

In the present paper, we will compare neural firing rate and interneural firing synchrony as potential representations of sound azimuth in AI. We show that codes based on neural synchrony generally provide sharper azimuth tuning than codes based on firing rate. If neurons in hierarchically higher cortical fields than AI perform an interneural synchrony detection operation, a more sharply tuned representation of sound azimuth may result. We also show that a population-vector model based on either peak firing rates or on coincident firings can extract sound azimuth in a rather coarse way. No difference in the performance of the predictor was observed for peak firing rate or neural synchrony, suggesting that the directional information that can be extracted by the population vector procedure from these measures is the same.

METHODS

The care and the use of animals reported on in this study was approved (No. P80905) by the Life and Environmental Sciences Animal Care Committee of the University of Calgary.

Animal preparation

Cats were premedicated with 0.25 ml/kg body weight of a mixture of 0.1 ml Acetromazine (0.25 mg/ml) and 0.9 ml of atropine methyl nitrate (0.5 mg/ml) subcutaneously. After ∼0.5 h, they received an intramuscular injection of 25 mg/kg of ketamine (100 mg/ml) and an intraperitoneal injection of 20 mg/kg of pentobarbital sodium (65 mg/ml). Lidocaine (20 mg/ml) was injected subcutaneously and rubbed in gently, then a skin flap was removed, and the skull cleared from overlying muscle tissue. A large screw was cemented upside down on the skull with dental acrylic. A 4- or 8-mm-diam hole was trephined over the right temporal cortex. The dura was left intact, and the brain covered with light mineral oil. Then the cat was placed in a sound-treated room on a vibration isolation frame (TMC micro-g), and the head secured with the screw. Additional acenpromazine/atropine mixture was administered every 2 h. Light anesthesia was maintained with intramuscular injections of 2–5 mg·kg⁻¹·h⁻¹ of ketamine. The wound margins were infused every 2 h with Durocain and also every 2 h new paraffin oil was added if needed. The temperature of the cat was maintained at 37°C. At the end of the experiment, the animals were killed with an overdose of pentobarbital sodium.

Acoustic stimulus presentation

Stimuli were presented with an array of nine loudspeakers (Realistic Minimus 3.5) placed in a semicircle with radius of 55 cm in the horizontal plane. The cat’s head was in the center of the semicircle, the speakers were separated by 22.5°. The speakers at −90°, −67.5°, and −45° will be referred to as contralateral, those at −22.5°, 0, and 22.5° as frontal, and those at 45, 67.5, and 90° as ipsilateral. Consequently, contralateral units were defined as those having a best azimuth (at threshold) or preferred azimuth (at 25 dB above threshold) for speakers at −90°, −67.5°, and −45°. Frontal units had a best or preferred azimuth at −22.5°, 0, and 22.5°. Ipsilateral units had best or preferred azimuths at 45, 67.5, and 90°. Omnidirectional units were defined as having a bandwidth at 25 dB above threshold of ∼180°. i.e., a 50% response criterion was obtained for all units. The azimuth with the maximum response was the preferred azimuth. Multipulsed units had more than one well-defined peak within the response area, the highest peak at 25 dB above the lowest threshold was used for the assignment of preferred azimuth.

The sound-treated room was made anechoic for frequencies >625 Hz by covering walls and ceiling and exposed parts of the vibration isolation frame and equipment with acoustic wedges (Sonex 3°). Calibration and monitoring of the sound field was done using a B&K (type 4134) microphone placed directly above the animal’s head, facing the loudspeakers. Random-frequency tonepips, noise burst, and clicks were used to locate units. CF and tuning curves of the isolated neurons were determined with 50-ms duration, gamma-shape envelope, tonepips presented randomly in frequency once per second (Eggermont 1996a).

After the CF was determined, 100-ms duration pseudorandom noise bursts with 5-ms linear rise-fall time were presented once per second from a randomly selected speaker. The pseudorandom noise itself was randomized further by drawing from 10 different seed numbers for the noise generation. For each azimuth, 50 noise bursts were presented, thus including five repetitions each of the 10 different random sequences, and the entire stimulus ensemble for one intensity level lasted 450 s. At each azimuth, the same set of 10 random noise bursts was presented. Intensity levels were presented interleaved usually starting at 75 dB SPL, then 55, 35, and 15 dB and subsequently at the intermediate intensities. Around threshold 5 dB steps were usually made.
Recording and spike separation procedure

Two tungsten microelectrodes (Micro Probe) with impedances between 1.5 and 2.5 MΩ were advanced independently perpendicularly to the AI surface using remotely controlled motorized hydraulic microdrives (Trent-Wells Mark III). The electrode signals were amplified using extracellular preamplifiers (Dagan 2400) and filtered between 200 Hz (Kemo VBF8, high-pass, 24 dB/oct) and 3 kHz (6 dB/oct, Dagan roll-off) to isolate spikes from local field potentials (LFPs). The signals were sampled through 12 bit A/D converters (Data Translation, DT 2752) into a PDP 11/53 microcomputer, together with a timing signal from two Schmitt triggers. In general the recorded signal on each electrode contained activity of more than one neural unit. The PDP was programmed to separate these MU spike trains into SU spike trains using a maximum variance algorithm (Eggermont 1991; Salganicoff et al. 1988). We added a feature that allowed us to save the waveforms both of the learning set and those during the actual experiment so that we could examine in retrospect the quality of the separation. In addition, an inclusion distance (in standard deviations) from the center of each cluster could be selected (usually 2 SD was chosen); spikes outside these areas were classified as class 0 and not stored. The spikes from the separation classes, each assumed to represent a particular neuron, were coded for display. The unit code plus the time of the spike occurrence were send to the MicroVax II, which presented on a Macintosh LC475 an on-line color-coded MU dot display organized per tone frequency or speaker position.

In addition the electrode signals were band-pass filtered (24 dB/oct, −96 dB/oct) between 10 and 100 Hz to obtain spike-free signals of ongoing LFPs. These signals also were passed through Schmitt-triggers set at about 2 SD below the mean value of the ongoing signal during silence. The ‘spikes’ produced by these signals were processed in the same way as SU spike data. We have shown previously (Eggermont and Smith 1995a) that the triggers produced by these level crossings represent most of the temporal response properties of the SUs recorded at the same electrode.

The boundaries of the primary auditory cortex for cases with the 8-mm hole were explored by taking a series of LFP and MU measures (with the high-pass filter set at 10 Hz with 6 dB/oct) from caudal to rostral and assuring that there was a gradual increase in CF, which after a region with no responses to tones (because of our limited high-frequency stimulus range) reversed in direction. These boundaries were indicated on a drawing of the cortical surface showing the location of the blood vessels and gyri. From this map, we estimated the desired CF location for our recordings and inserted the electrodes in that location. In the 4-mm-hole animals, the simultaneous recording with two electrodes had to satisfy the criterion that the more anterior electrode showed the highest characteristic frequency. Recordings were generally made in the dorsal region of AI and between 300 and 1,000 µm below the cortex surface.

Spike-separation procedure

The single-electrode recordings in this experiment always consisted of spikes from more than one unit. Using spike separators to sort out the SU spike trains from a MU record has several inherent weaknesses (Gray et al. 1995). One is that superimposed spikes are sorted in a separate class; such spikes thus are removed effectively from the SU spike trains and thus from the analysis, unless one resorts to very time-consuming decomposition techniques. Because the average percentage of noise bursts that produces a short-latency spike in a SU at the optimum azimuth in our recordings was at most 50% (comparable with the value found for moderate to high-intensity clicks) (Eggermont 1991), and assuming that a second unit fires independently from the other, the percentage of noise bursts that produces overlapping spikes, i.e., fire within 2 ms, is <10%. At low sound levels, the variability in latency is much larger so the probability of overlap will reduce correspondingly. We do not expect that for our recording condition overlap occurring in a SU spike recording is significant in SU records that are different from well-isolated SU recordings.

Another factor that could influence the results is misclassification of spikes. Our off-line quality-check procedure, based on a print out of a randomly selected 10% of the sorted spike waveforms per electrode, allowed 5% misclassified spikes per class to be acceptable. Thus if 1,000 spikes were recorded and classified as a particular SU, ~100 would be printed out and finding more than five distinctly different traces would be sufficient to reject the sorted spike train as SU. Classes with >5% of misclassifications were not further considered. This eliminated about one-third of the potential number of SUs in our study. More serious problems may occur when spikes in a burst show a decrease in amplitude and the smaller ones tend to be classified as a separate neuron. This would likely have the largest effect at low intensities where burst lengths are largest. However, because spike bursts in our recordings were rarely longer than four spikes (cf. Bowman et al. 1995), we have not found any evidence for this in our data. In summary, we have no reason to assume that omissions and misclassifications of spikes are responsible for such differences in response properties as finding SUs with ipsilateral and contralateral preferences at 25 dB above threshold in recordings on the same electrode.

Data analysis

The action potentials in a 100-ms time window after the noise-burst onsets were counted for each intensity and speaker location. Because the noise bursts lasted 100 ms, off responses were outside the analysis window and not incorporated in the analysis. The results per stimulus intensity were combined into a count-location-intensity profile from which iso-count contours as a function of azimuth, count-intensity functions for each azimuth and iso-intensity-count contours could be derived. The azimuth tuning curve was an iso-count contour as a function of azimuth defined for a count equal to 30% of the maximum spike or LFP trigger count across all intensities and stimulus azimuths.

PSTHs were made for each speaker location and covered 100 ms after stimulus onset in 1-ms bins. PSTHs were smoothed with a rectangular 5-bin window to facilitate peak latency readings. Peak count amplitudes also were measured from the smoothed PSTHs.

Threshold levels were estimated from the PSTHs. If 10-db steps were used, the threshold was assigned at 5 dB below the lowest level that showed visible locking. If 5-db steps, were used the threshold level assigned was 2.5 dB below the last active one. Best azimuth was defined as that azimuth that resulted in the lowest threshold.

Preferred azimuths were estimated at 25 dB above the lowest threshold and were defined as the midpoint of the range where the response parameter in question (overall count or PSTH-peak count, peak coincidence count) was 50% of maximum over all intensities and stimulus azimuths (i.e., within the 50% contour line of the response surface). The azimuth tuning bandwidth was estimated at 25 dB above the lowest threshold as the angle over which the response measure was within the 50% contour.

The peak coincidence count used in most analyses is taken as the peak value of $R_{ab}(\tau)$, the cross-correlogram of spike trains A and B. In some analyses, the neural synchrony coefficient, defined $R(\tau) = \frac{R_{ab}(\tau) - E}{\sqrt{N_A N_B}}$, where $E = N_A N_B/N$ is the expected value under the assumption of independent firings in the two spike trains. $N$ is the number of 1-ms bins and $N_A$ and $N_B$ are the numbers of spikes in the A and B trains in the set of 100-ms windows of the record (Eggermont 1994).

Synchronization codes for azimuth representation were explored
by calculating cross-coincidence histograms for 5-ms bins for SU activity and LFP triggers on the same electrode, between SU activity on the same or both electrodes, and between SU and MU activity on one electrode with MU activity on the other electrode. All these analyses were done for each azimuth and each intensity so that again intensity-location response surfaces could be constructed. The response parameter plotted in these cases is the peak coincidence count, and the plots are referred to as peak coincidence surfaces.

**Population vector**

The basic assumptions underlying a population vector code (Georgopoulos et al. 1986) are that short-term firing rate is the relevant parameter, that firings of one cell are independent of those from another, and that each cell has a preferred azimuth but is otherwise so broadly tuned that it responds to sound from all azimuths. Preferred azimuths are assumed to be uniformly distributed across all possible directions. Typically the directional tuning curves are cosine shaped

\[ r_i(\theta) = b_i + c_i \cos(\theta - \theta_0) \]  

(1)

where \( r_i(\theta) \) is the firing rate of unit \( i \) for sound from azimuth \( \theta \), \( b_i \) is the average firing rate over all sound directions, \( c_i \) is a constant, and \( \theta_0 \) is the preferred azimuth of the unit at that level. Each azimuth-sensitive unit makes a vectorial contribution along the preferred direction \( \theta_0 \) with a weight equal to \([r_i(\theta) - b_i] / c_i\) for sound from the direction \( \theta \). The population vector results from a vector addition of the individual unit vectors

\[ \mathbf{O}_{\text{est}} = \mathbf{R} \mathbf{O}_0 \]  

(2)

where \( \mathbf{O}_{\text{est}} \) is the population vector containing the strengths and direction of the population response for sound from azimuth \( \theta \), \( \mathbf{R} \) is the matrix containing as elements \( r_i(\theta) \) the relative firing rates \([r_i(\theta) - b_i] \) and \( \mathbf{O}_0 \) is the vector containing the best azimuths \( \theta_0 \) for all units \( i = 1, N \) in the form \([\cos\theta_0 + (-1)^j \sin\theta_0]\).

We assigned to each unit a best azimuth selected from the nine possible sound azimuths closest to the \( \theta_0 \) value resulting from a curve fit using Eq. 1. Subsequently, the firing rates of all neurons that had the same best azimuth were averaged for each of the nine sound presentations. Consequently the calculation of the population vector \( \mathbf{O}_{\text{est}} \) was based on nine best-azimuth groups for nine sound azimuths, making \( \mathbf{R} \) a \( 9 \times 9 \) matrix. The averaging action removed the preference bias for contralateral best azimuths in the sample of units encountered in AI and ensured a uniform distribution. A normalization based on the subtraction of the mean firing rate of each unit across sound azimuths reduced the variability in the firing rates entered into \( \mathbf{R} \) and assured that each unit contributed about equally to the estimate of the mean peak firing rate. Under these conditions one expects the population-vector model to show its best performance.

All calculations were performed using MATLAB 5 on a Macintosh Power PC. Statistical analyses were performed using Statview 4.5 and additional data plotting was done with Horizon or PowerPoint software.

**Results**

Recordings were obtained in 13 juvenile ( \( >33 \) days) and 13 adult cats ( \( >70 \) days). In only five of the juvenile cats was a complete intensity series obtained so the data from 58 MU recordings and 36 LFP recordings, comprising 102 isolated SUs, presented here are for 18 cats. For each of these recordings, noise bursts were presented 50 times each for 9 azimuth directions and on average 7 intensities. Except for minor differences in latencies and absolute threshold values (Eggermont 1996b), the two age groups showed similar azimuth dependent results. The MU data used in the following analyses were restricted to the combination of isolated SUs recorded on that electrode; unclassified spikes were not included.

**Example for firing rate measures**

Figure 1 shows MU dot displays (Fig. 1, A and B) and LFP-trigger displays (Fig. 1C) for simultaneous recordings on two electrodes in a 86-day-old cat at a depth of 700 \( \mu \)m below the surface in the 7-kHz iso-frequency strip \( \sim500 \) \( \mu \)m apart in the dorsoventral direction. Only LFP triggers for electrode 1 are shown, those for electrode 2 were qualitatively the same. Both MU recordings consisted of three isolated SUs, identified by different colors, with similar response properties. The intensity range used was from 5 to 60 dB SPL. We show the results from 10 to 55 dB SPL. At the lowest stimulation levels notable spontaneous activity is seen. High-stimulus levels (\( >45 \) dB SPL) produced a nearly complete suppression of spontaneous activity after the 20-ms-duration on-response. The LFP triggers at levels \( >25 \) dB SPL show secondary responses at 50- to 60-ms latency that do not have a correlate in the MU response (Fig. 1, A and B), whereas the initial LFP triggers correlate very well with the onset spike activity. At high-stimulus levels, no obvious preference for stimulus azimuth is seen, albeit that there are small systematic latency differences, which is the topic of the companion paper (Eggermont 1998). At levels \( <30 \) dB SPL, the response to sound from the frontal speaker locations is longer in latency and less robust; this holds both for MUs and LFP triggers. Note that the MU-response duration at near-threshold values, which are 20 dB for the frontal speaker and 10 dB for the contralateral speakers, is considerably prolonged compared with that at suprathreshold levels. This tendency of longer-duration bursting close to threshold values was a general finding for the type of stimulus we used. We have noticed this before for tone-burst stimulation (Eggermont and Smith 1996b), and it is likely the result of a reduced postactivation suppression at lower intensities (Phillips and Sark 1991).

An example showing SU responses with contralateral, frontal and ipsilateral azimuth preference is presented in Fig. 1 in the companion paper (Eggermont 1998). Unit 1,3 is a typical contralateral unit, unit 2,2 is a frontal unit, and unit 2,3 is an ipsilateral unit.

For the recordings shown in Fig. 1 of the present paper, Fig. 2, A–C, shows the intensity-azimuth dependence of the total number of spikes in 100 ms (overall count) after noise-burst onset and Fig. 2, D–F, shows the intensity-azimuth dependence for the PSTH peak counts in the same time window. The contour plot levels are 15% (no shading), 30% (light shading), 50% (medium shading), and 70% (dark shading) of the maximum response for each plot. The global features of the azimuth representation are similar for the overall spike count (Fig. 2, A–C) and for the peak count of the MU PSTHs (Fig. 2, D and E). For the recordings on electrode 1 (Fig. 2D), a peak response <70% of maximum was found over a larger intensity range than for the overall spike count (Fig. 2A). The PSTH-peak response for the activity on electrode 2 (Fig. 2E) was found at a slightly
higher intensity level than that for the mean rate. Note that both MU records tend to have nonmonotonic rate-intensity functions especially for contralateral sound positions. The LFP-trigger representation shows that the peak amplitude is saturating >30 dB SPL, whereas the overall trigger count is monotonically increasing throughout the entire azimuth range largely because of the secondary responses. The MU and LFP obviously capture different aspects of the cortical neural activity.

**SU and MU synchrony measures**

For a recording from the same electrode tracks as the previous examples, Fig. 3A shows the shape of the cross-coincidence histograms between SU activity on each of the two electrodes as a function of azimuth for four intensity levels. The peak coincidence count is found most frequently at a lag of 0–10 ms, suggesting that on average the latencies of the firings on electrode 2 are 5–10 ms longer than those on electrode 1. Figure 3B shows the peak coincidence surface for this recording; one notices that the preferred azimuths again are in the contralateral field. The coincidence histograms in Fig. 3A are relatively broad and indicative of rate covariance rather than event correlation (Eggermont and Smith 1995b). This was further explored by attempting to predict the peak coincidence surface from multiplication of the “overall spike count” surfaces for the two SU recordings. Under the assumption of independence of firing for the two units, the expected peak value of the cross-coincidence histogram, independent of the lag time, is equal to \( N_A N_B / N \) (Eggermont 1992), where \( N_A \) and \( N_B \) are the number of firings of the two units and \( N \) is the total number of bins in the 450 analysis windows, each of 100-ms duration, after stimulus onset. Because the number of bins only provides a scaling, we cross-multiplied the number of spikes at each azimuth-intensity combination to obtain the expected shape of the coincidence surface. This multiplication surface (Fig. 3C) predicts the shape of the coincidence surface qualitatively, suggesting that rate covariance is a dominant contributor to the correlation. Note, that normalizing coincidence counts for overall firing rate could be done by dividing by the length of the record (number of stimuli \( \times 100 \) ms) per azimuth and intensity level. This will not change the appearance of the surface plots and leave the relative scaling intact. Converting the original and predicted peak coincidence plot to synchronization coefficient plots shows that the predicted peak values are a factor 4 smaller than that for the simultaneous synchrony, suggesting that event correlation in this case plays an important role.

Figure 4 shows examples of several neural synchrony measures, as a function of azimuth and intensity for the same data set shown in Figs. 1 and 2, displayed as contour plots for levels of 15% (not shaded), 30% (light shading), 50% (medium shading), and 70% (dark shading) of the maximum response. Figure 4A uses the peak number of coincidences between SU firings (green in Fig. 1A) and the LFP triggers on electrode 1. For intensities <35 dB, the highest number of coincidences is found for contralateral speaker positions; at higher intensities, the peak number of coincidences spreads to include frontal and ipsilateral directions. Figure 4B illustrates the findings for the peak number of coincidences of the firings of a SU (blue in Fig. 1B) from electrode 2 with the LFP on that electrode. Best azimuths between 25 and 45 dB are found for contralateral speaker positions. Figure 4C shows the azimuth-intensity dependence for the overall spike count for the green SU from the MU record shown in Fig. 1A. This allows a comparison of the azimuth dependence overall spike count and coincident spike counts for a SU. Figure 4D shows the peak coincidence count for the firings of two SUs (green and blue) from Fig.
Fig. 2. Overall-firing rate and peak-firing rate in 100 ms after stimulus onset as a function of sound azimuth and stimulus intensity for the 3 conditions shown in Fig. 1. A and B: contour plots for 15, 30, 50, and 70% of maximum overall MU firing rate for electrodes 1 and 2. C: same for LFP triggers. D and E: results for peak MU firing rate from the poststimulus time histograms (PSTHs) for electrodes 1 and 2. F: same as D and E except for the LFP triggers. Filled arrows indicate the position of a spatial ‘‘null,’’ and the open arrows the best azimuths.

1A. Figure 4E shows the peak coincidence count for SU activity from electrode 1 (green) with MU activity from the other electrode. Figure 4F shows the peak coincidence count for MU activity from the two separate recordings shown in Fig. 1, A and B. The preferred azimuth for the peak coincidence count of the firings of the two units on electrode 1 is in the contralateral field (Fig. 4D), and not surprisingly the peak coincidence count for the firings of the SU on electrode 1 with the MU activity on electrode 2 (Fig. 4, E and F) also shows a peak in the contralateral field. A comparison between Fig. 4, E and F, suggests that the peak coincidence count for MU by MU activity essentially captures the peak coincidence between SU but with better statistics because of the larger number of coincidences. The best azimuth for Fig. 4, both F and E, is at −67.5°, but the response area is much more restricted in case of the peak coincidence count for the MU recordings.

**Azimuth preference at 25 dB above threshold**

For a criterion level of ≈50% of the maximum response, 28/36 of the LFP recordings had an omnidirectional sensitivity at 25 dB above threshold (as in the example in Figs. 1C and 2C) and 8/36 had a contralateral sensitivity. By contrast, the MU recordings showed only 10 omnidirectional response preferences, and in 17/58 cases, a contralateral preference. 9 recordings had a preference for the ipsilateral field, 9 recordings were found with frontal sensitivity, and 13 recordings with multidirectional preferences. At threshold values all 36/36 LFP recordings showed preferences for contralateral speaker sites, i.e., those located at −45, −67.5, and −90°. The MU recordings showed in 26/58 cases a preference for the −67.5° speaker and in 17 other cases for the neighboring speakers (−45 and −90°). Nine cases showed also shows a peak in the contralateral field. A comparison between Fig. 4, E and F, suggests that the peak coincidence count for MU by MU activity essentially captures the peak coincidence between SU but with better statistics because of the larger number of coincidences. The best azimuth for Fig. 4, both F and E, is at −67.5°, but the response area is much more restricted in case of the peak coincidence count for the MU recordings.

For MU recordings for which a preferred azimuth at 25 dB above threshold could be assigned, the majority of the individual SUs generally had the same azimuth preference as the MU activity on that electrode but with a larger proportion of ipsilateral best azimuths. This is illustrated in the scatterplot (Fig. 5) that shows a strong correlation between best azimuths for SU and MUs for the majority of recordings, albeit that there are 19 SUs that had ipsilateral preferences, whereas the corresponding 12 MU records showed contralateral preferences. In all of these 19 cases there were SUs on the same electrode that had contralateral preference and considerably higher firing rates than the ipsilateral unit. The
mean difference between the SU best azimuth and that for the corresponding MUs was 22.5° (1 speaker position) toward ipsilateral (paired t-test, \( P < 0.0001 \)).

Behavior of the spatial ‘null’

An azimuth value without a response at a particular intensity level flanked by azimuths producing clear responses, which we call a spatial “null,” was observed in 21 MU recordings in adult cats only. It is visible in the dot displays (Fig. 1A) and contour plots (Fig. 2, A and D; indicated by the arrows) at \( \approx 45° \) in the ipsilateral field. The level at which the null appeared (20 dB SPL in this example) was related closely to the threshold level for best azimuth. This is illustrated in Fig. 6A, which shows a scatterplot of the highest intensity at which the spatial null was present versus best azimuth. The regression line has a slope not significantly different from one. The mean difference between the appearance of the null and best azimuth threshold was 11 dB. It is notable that nulls only appeared in recordings with best azimuths in the contralateral field. The most frequent null was found for a speaker direction in the frontal field, 22.5° ipsilateral to the midline (Fig. 6B). As can be observed from the dot displays in Fig. 1, the appearance of the spatial null is quite abrupt, and for stimuli at those azimuths, the response latency increases dramatically combined with a more variable firing pattern.

Dependencies on CF

The range of CF values in this study was from 3 to 25 kHz. The threshold level for best azimuth was independent of CF, and the distribution of threshold values was quite broad as previously found also for toneburst stimulation (Eggermont 1996a). The average best-azimuth threshold was 18.7 ± 9.8 (SD) dB SPL. The best azimuth was also independent of CF with a modal position in the contralateral field at \(-67.5°\). The position of the spatial null was also independent of CF. The various bandwidth measures listed in Table 1 were all independent of CF (regression analysis, all \( r^2 \) values were <0.006).

Neural synchrony versus firing rate representations

As we have seen in the example recording (Figs. 1–4), there was no qualitative difference in azimuth representation based on overall firing rate, peak firing rate, and synchrony measures obtained from the various cross-correlation calcu-
FIG. 4. Firing rate and firing synchrony measures as a function of sound azimuth and stimulus intensity displayed as contour plots for 15, 30, 50, and 70% of maximum response. A: peak coincidence counts for a SU (green) from the MU record of Fig. 1A with the LFP triggers shown in Fig. 1C. B: peak coincidence counts for a SU (blue) from the MU record of Fig. 1B with the LFP triggers from that electrode. C: result for the overall spike count for the ”green” SU from the MU record shown in Fig. 1A. D: peak synchrony for the firings of 2 SUs (green and blue) from Fig. 1A. E: peak synchrony for SU activity from electrode 1 (green) with MU activity from the other electrode. F: peak synchrony for MU activity from the 2 separate electrodes shown in Fig. 1. Arrows indicate the point where the 50% bandwidth measures are taken, the values are 60 and 50°, respectively.

The sound azimuths for which several response measures were maximal at 25 dB above best azimuth threshold correlated generally with best azimuth. This is shown in Fig. 7B for the azimuth showing the maximum cross-correlation between the MU firings on the two electrodes with MU best azimuth ($r^2 = 0.17$, $P = 0.008$) and in Fig. 7C for the azimuth with the maximum peak firing rate and MU best azimuth ($r^2 = 0.11$, $P = 0.029$). For the sound azimuth showing the largest coincidence between the SU firings and LFP triggers at 25 dB above threshold, no significant correlation with best azimuth for LFP was found (Fig. 7D). A large group of recordings was found with ipsilateral preferences at higher levels but best azimuths for contralateral speaker-positions at threshold.

A comparison between azimuth preferences based on firing rate or interneural synchrony measures at 25 dB above threshold is given in Fig. 9. Figure 8A compares the preferred azimuth for highest MU peak firing rate with preferred azimuth for the highest MU overall firing rate; as expected a strong correlation ($r^2 = 0.45$, $P < 0.0001$) is found. Figure 8, B and C, compares the preferred azimuths for the peak values of two coincident count measures with the preferred azimuth for the highest MU overall firing rate. The correlation is stronger for the peak coincident counts of SU firings with LFP triggers ($r^2 = 0.47$, $P < 0.0001$) than for the peak coincidence count between MU firings on the two electrodes ($r^2 = 0.22$, $P = 0.001$). Figure 8D shows that the azimuth for which the largest peak coincidence between SU firings and LFP triggers is found also is correlated strongly with that for the MU peak firing rate on the same electrode ($r^2 = 0.69$, $P < 0.0001$).

Figure 9A shows the 50% contour bandwidth at 25 dB above threshold for the overall SU spike count compared with that for the MU spike count on the same electrode. The SU 50% contour bandwidths are significantly narrower ($P = 0.03$) than those for the MUs, largely because of the
The mean and standard deviations of the various bandwidths are shown in Table 1. For the complete data set, pairwise comparison of the three synchrony measures showed that the LFP-synchrony measure was significantly narrower than both the stimulus-synchronous peak-count measure and the dual-electrode MU-synchrony (which were not significantly different). If omnidirectional MU recordings are excluded, the neural synchrony bandwidths are significantly smaller than the stimulus synchrony (expressed in the peak PSTH values) bandwidth. This suggests that interneural firing synchrony displays a sharper azimuth tuning than does stimulus synchrony.

Offset and "late" responses

Late responses and OFF responses generally occurred near the end of the 100-ms noise burst, and because they behaved generally differently to sound, the following definition is presented. OFF responses follow the termination of the 100-
TABLE 1.  **Directional tuning bandwidth at 25 dB above best azimuth threshold**

<table>
<thead>
<tr>
<th>Bandwidth</th>
<th>No omnidirectional units, deg</th>
</tr>
</thead>
<tbody>
<tr>
<td>All data, deg</td>
<td></td>
</tr>
<tr>
<td>100-ms count</td>
<td></td>
</tr>
<tr>
<td>Single unit</td>
<td>112 ± 50</td>
</tr>
<tr>
<td>Multiunit</td>
<td>117 ± 58</td>
</tr>
<tr>
<td>Peak count</td>
<td></td>
</tr>
<tr>
<td>Multiunit</td>
<td>85 ± 41</td>
</tr>
<tr>
<td>Peak coincidence count</td>
<td></td>
</tr>
<tr>
<td>SU \times LFP</td>
<td>72 ± 29</td>
</tr>
<tr>
<td>MU \times MU</td>
<td>86 ± 58</td>
</tr>
</tbody>
</table>

Values are means ± SD. SU, single unit; MU, multiunit; LFP, local field potential.

ms noise burst with a latency of ≈15 ms. Late responses are those that occur either earlier than the OFF responses or those that occur with a latency of >50 ms after the end of the noise burst. Of the 58 recordings, 14 only had onset responses, in one case there was only an OFF response, and in 2 cases there was only a late response. In 20 recordings both an onset and OFF response were present, in 14 recordings the onset was accompanied by a late response, and in 5 recordings there was an onset, an OFF response, and a late response. The majority of the three types of response showed similar best azimuth; however, in three cases, the OFF response was differently tuned from the onset response and in four cases, the late response was differently tuned from the onset response. For these cases, the OFF responses were more broadly tuned than the onset responses, whereas the late responses often showed a tuning that was complementary to the onset response. Thus if the ON response was tuned contralaterally, as they mostly were, then the late response in these cases was tuned ipsilaterally.

**Population vector predictions**

It is unlikely that sound localization is based on the activity of SU or MU recordings from one site, but it is plausible that a more global population code exits. Population codes combine for instance the firing rate of all neurons involved and combine them into a population vector. Each neuron is entered with its firing rate as a function of sound azimuth. This vector does contain its preferred azimuth. In this global coding strategy, for a sound presentation from a given azimuth, each neuron votes for its preferred azimuth with its share, its firing rate. The weighted sum of all the votes results in an estimated azimuth. By analyzing this strategy, we may be able to find how much information is present in the population activity and also whether synchrony measures contain more information than rate measures. Ultimately we may be able to say whether such a strategy is likely to be employed by the auditory cortex. The population decoding analyses in the present study are based on the 102 isolated SUs. The
Correlations between azimuth-tuning curve bandwidth measures at 25 dB above best azimuth threshold. A: scattergram for the bandwidth estimates for single unit spike count in 100 ms and the MU spike count in the same interval. B: comparison of the bandwidths for MU PSTH peak count and for overall count in 100 ms. C: comparison of the bandwidths for the peak coincidence count between MU spikes and LFP triggers and the MU spike count in 100 ms. D: comparison of bandwidths for the number of coincident spikes in the 2 MU records and that based on MU spike count in 100 ms.

Predictions on basis of peak firing rate were carried out separately for the low-intensity group for which the directionality of the pinna limits the response largely to the contralateral field and for the high-intensity group where responses for sounds from all azimuths are found. We also explored whether the peak coincidence count, reflecting the maximal synchrony between two SUs or between SU and local field potential, would present different information from spike count. For cross-correlations of the 102 SUs with the corresponding LFP triggers, the peak cross-coincidence histogram values at each azimuth and intensity level were found to be quite good, and the predicted azimuths from the fit correspond, with a few exceptions, with the assigned best azimuths. Here one has to take into account that the preassigned azimuths had to be one of the nine speaker azimuths. Units tuned to contralateral speaker positions on average responded with much higher peak firing rates than units tuned to ipsilateral azimuths. Subtracting the mean firing rate across all sound azimuths for each best azimuth group results in a change-in-spike count profile (Fig. 11, for the high-intensity group). This illustrates that the largest differences from mean spike counts are found at sound azimuths corresponding to best azimuth, i.e., along the equal azimuth diagonal. The contour lines shown are in fractions of the maximum and ~0.1 apart. Dark shading corresponds to smaller values.

Summed PSTHs across all individual units are shown in Fig. 10 for intensities of 15–65 dB SPL. All histograms on the same scale. At 15 dB SPL, clear summed stimulus-locked activity is only found for contralateral sound azimuths, namely between ~45° and ~90° from the midline. At 25 dB SPL, ipsilateral sound also evokes stimulus locked activity, whereas the frontal speakers are less effective. For higher stimulus intensities, the responses start to show complex latency and peak firing rate dependencies on azimuth, especially for frontal sound sources. This variability in peak
calculated and averaged across units with the same best azimuth for the two intensity groups. The average number of peak coincidences for the high-intensity group ranged from 2.33 (best azimuth 22.5°, sound azimuth −90°) to 13.5 (best azimuth and sound azimuth at −67.5°). For the low-intensity group, the average number of peak coincidences was somewhat lower and ranged from 1.5 (best azimuth 45°, sound azimuth 0°) to 10.53 (best azimuth and sound azimuth at −90°). For both intensity groups, the highest number of coincidences is found for units with contralateral best azimuths, which also showed the largest firing rates. The mean number of coincident counts for each sound azimuth was subtracted before the predictions were made. We show the predicted azimuths compared with the actual stimulus azimuths for peak firing rate for the two intensity groups and for peak synchrony at the high-intensity group only (Fig. 12). One observes that the predictions based on rate and synchrony virtually overlap for the high intensities. For the low-intensity, group synchrony measures were only available in the contralateral field because of the restricted responses of the LFPs. Meaningful predictions based on this limited synchrony representation were not possible.

We also compared the predictions for best azimuth based on groups defined according to their best azimuth (i.e., at threshold) and on basis of the preferred azimuth. The predictions on basis of the two azimuth assignments were similar in shape but the threshold azimuth assignment was representative for that shown for the lower intensities, and the preferred-response assignment was very similar to that used in the analysis shown in Fig. 12. Thus assigning different azimuths keeps the overall shape of the prediction curve the same but shifts it along the sound azimuth axis.

These combined data suggest that there is more directional information in the population activity at high intensities than at low intensities. At low intensities, the information allows a decision between contralateral and ipsilateral azimuths with a strong bias to ipsilateral. At high intensities, this bias is reduced to 20° for midline azimuth. For sound presented from azimuths of ±22.5°, the predicted azimuths were about ±65°. The prediction suggests that, for this data set, the information, extracted by the population vector method at intensities >40 dB SPL, from peak neural synchrony measures, is not different from that present in peak firing rates. This is despite the fact that azimuth tuning curves are narrower for neural synchrony than for peak firing rate.

**Discussion**

We have shown that there is no qualitative difference between azimuth preferences based on firing rate, stimulus synchrony, or neural synchrony and that most recordings showed a maximum for all these measures for sound presented in the contralateral hemifield. A spatial null direction was found in adult cats for recording sites with CFs between 3 and 25 kHz and at stimulus levels that were on average 11 dB above best azimuth threshold. Quantitatively, the neural synchrony measures based on coincident spike counts were more sharply tuned to azimuth than a stimulus-synchrony measure based on the peak count in the PST histogram. Interneural synchrony measures were also more sharply tuned than overall spike-count measures in a time window of 100 ms after stimulus onset. If neurons in hierarchically higher cortical fields would perform a coincidence or co-firing detection operation, a more sharply tuned representation of sound azimuth may result. Azimuth predictions based on a population vector model suggest that the amount of information on sound azimuth present in peak PSTH count and peak coincidence count, as extracted by this computational coding procedure, is the same. The relatively large error in the predictions suggests that, despite the coarse grain of the speaker positions, the observed behavioral orienting accuracy of 4–5° around the midline cannot be based on the range of any of the measures of AI activity evaluated in this study.

**Comparison with other studies**

The distribution of azimuthal sensitivity based on overall MU firing rate is largely similar to those reported in the literature. At a level of 25 dB above threshold, we found that 17% of our recordings were omnidirectional; this is similar to the 17.6% reported by Rajan et al. (1990a). About 16% of our recordings showed a sensitivity for sound from frontal speakers; this is comparable with the 17% that was reported by Imig et al. (1990) and Rajan et al. (1990) but larger than the 7% found by Bruggé et al. (1996). Again, about 16% of our recordings were sensitive to ipsilateral sound; this is close to the 20% found by Rajan et al. (1990a) and the 23% reported by Imig et al. (1990). Differences exist in the numbers for contralateral sensitive units: whereas we find only 29%, the values in the literature range from 43% (Rajan et al. 1990a) to 60% (Imig et al. 1990). The difference appears to be made up by a potentially different interpretation of multipeaked sensitivities for which we report 22%, whereas it was only ~6% in Rajan et al. (1990a) and none in Imig et al. (1990). Within 10 dB from threshold, 74% of our recordings showed a best azimuth for the contralateral field, 16% for frontal locations, 5% remained omnidirectional, and 5% preferred ipsilateral locations.

The bandwidth at 25 dB above threshold averaged across
all “tuned” MU-recordings in our data using spike count as our criterion was 77°, whereas when the omnidirectional units were included, it increased to 117°. This is comparable with results obtained by Clarey et al. (1995) averaged over all stimulus levels that gave 74°, both in AI and in MGB.

It has been shown (Clarey et al. 1994) that for MU and SU recordings, measures of azimuth preference and selectivity showed significant correlations. This is similar to what we have found: albeit that 19/102 SUs were ipsilaterally tuned, whereas the corresponding 12 MU recordings were contralaterally tuned. Another exception is the finding of more omnidirectional sensitivity among MU recordings. This potentially could result from a combination of opposite hemifield sensitivities for the constituting single neurons. However, analyzing the 10 omnidirectional MU recordings showed that only in one case one of the units was contralaterally and the other ipsilaterally tuned. In three cases, one of the individual SUs also had an 180° bandwidth and another three had large bandwidths so that the sensitivity for sound extended in both hemifields. In the remaining three cases, the directional sensitivity of the constituting SUs was in the ipsilateral field. The apparent discrepancy appears to relate to slightly different (by one speaker location or so) directional preferences. As a result, the summed activity appeared more broadly tuned than for the individual units.

Spatial null

We found these nulls, azimuths at which sounds did not produce a response at a particular intensity level and flanked by azimuths at which the same sound produced clear responses, for a large CF range (from 3 to 25 kHz). In our data the spatial null occurs at levels between 20 and 70 dB SPL, but always within about 11 dB from threshold. Because the null azimuths were on average slightly ipsilateral from the midline, they could be the result of head shadow effects for high-frequency components in the noise. For these azimuths, the sound shadow may cause the stimulus level at both ears to be below threshold. The fact that the nulls also were found for CFs around 3 kHz and at relatively low levels excludes this as the only contributing factor.

Comparison of LFP data and MU recordings

LFPs are considered to reflect the averaged extracellular field representing the common synaptic potentials of a large group of neurons in proximity to the microelectrode. Spike-triggered averages of LFPs in visual cortex (Eckhorn and Obermuller 1993), somatosensory cortex (Morin and Steriade 1981), and auditory cortex (Eggermont and Smith 1995a) showed that spikes generally occurred synchronous with the negative deflections of the LFPs recorded on the same or not too distant electrodes and that the spike-triggered average LFP had a characteristic triphasic shape. Suppressed firing was associated with the positive deflections of the LFP. We observed that none of the azimuth preferences as reflected in LFP triggers were in the ipsilateral field, whereas MU activity recorded on the same electrode could show a preference for the ipsilateral hemifield. Our spatial receptive fields for LFPs were based on triggers on the depth-negative-going wave, and, as a consequence, only excitatory activity is reflected in the triggers. Using a current-source density (CSD) analysis, Mitzdorf (1985) has shown that the dominant component of cellular activity in cortex that contributes to the CSD and LFP is the excitatory postsynaptic potential (EPSP). The LFP likely consists of the CSDs integrated over a volume of ~1 mm³. With synchronous activation of many cells in that volume, one expects the LFP to be proportional to the time derivative of the EPSPs (see Eggermont and Smith 1995a for more details). Thus inhibitory postsynaptic potentials (IPSPs) may not be reflected in the LFPs, especially as IPSPs will occur nearly simultaneously with EPSPs (Douglas et al. 1995) and have a slower time course, although they will influence the firing of pyramidal cells. From comparisons for LFPs and SUs in AI (Eggermont 1996a), it is clear that frequency tuning for SU firings can be, but does not have to be, much narrower than for LFP triggers.

Comparison of SU and MU results

The finding that in 12/58 MU recordings both ipsilaterally and contralaterally tuned SUs are present simultaneously in the same MU record appears to be at variance with findings that units with different azimuth-sensitive response properties are segregated spatially (Middlebrooks and Pettigrew 1981; Rajan et al. 1990b) but agrees with the finding that MUs and SUs often showed different suprathreshold response properties (Clarey et al. 1994). It is of course possible that in ~20% of our recordings the electrode was registering border activity from segregated contralateral and ipsilateral “columns,” but the likelihood of that happening is small given that with tungsten electrodes a SU typically can be held over only 100 µm. In 9/12 of these recordings, the cats were >75 days, and in those cases the MU bandwidths of these recordings were >85°. Recording depths for these nonhomogeneous MU recordings was between 440 and 990 µm, with an average of 640 µm, so likely all from layers III and IV.

Neural synchrony coding versus firing rate coding

We called the use of the spike count in a 100-ms window after stimulus onset “rate coding.” We relate the use of the peak count of the PSTH to a “stimulus-synchrony” code. We further equate the use of peak measures of synchronous activity between neurons at different electrodes or between neurons and LFPs at the same electrode with a “neural-synchrony” code. For transient stimuli, the firings of different neurons tend to become synchronized because each neuron responds in time-locked fashion to the stimulus. Thus one could argue that for conditions as in our experiment it is rather pointless to make the distinction among neural synchrony, stimulus synchrony, and firing rate. Because of the strong postactivation suppression at intensity levels of 15 dB and more above threshold, peak PSTH count and overall spike count in a 100-ms window generally closely are related at suprathreshold levels so that a distinction between stimulus synchrony and firing rate is generally small. However, we have shown that there is in fact a quantitative difference in azimuth-tuning curve bandwidth for the various spike-count and spike-synchrony representations that we considered.
It is also undoubtedly true that stimulus-synchrony contributes to neural synchrony and sometimes may be the dominant factor especially in cases where the coincidences in firing result from a covariance in firing rate between the units. In this case, the product of the spike counts in a 100-ms window for two neurons is a good predictor for the peak neural synchrony between them as it theoretically should be (Eggermont 1992; Gochin et al. 1989). From Table 1 it is clear that the neural synchrony bandwidth at 25 dB above threshold is on average ~60% of the bandwidth for a firing-rate measure or a stimulus-synchrony measure. Thus neural synchrony, especially between the activity on different electrodes, provides a greater selectivity for stimulus azimuth than the SU or MU firing rate. This may be advantageous if converging activity from AI cells is used by coincidence detecting neurons in higher cortical areas such as the AES where cell firings may represent a “panoramic” code of azimuth (Middlebrooks et al. 1994). A similar finding was reported in cat visual cortex (VI) where the neural-synchrony receptive-field width was significantly smaller, ~20%, than that for the individual constituent single cells (Ghose et al. 1994). This reduction is comparable with the 40% found between SU bandwidth and the synchrony measures in our data.

The neural synchrony between SUs and similarly between MU activity on separate electrodes in all cases could be predicted qualitatively from the firing rates in the 100-ms time window after stimulus onset by simple cross-multiplication. In addition, predictions on the basis of the peak firing rates from the PSTHs were qualitatively similar to those based on overall firing rate. However, these were substantially smaller in size. Thus although stimulus synchrony is part of the overall synchrony, the trial-by-trial variation in the firing rates of the units is quite large, and the neural synchrony based on the within trial covariation in firing rate tends to be larger than the across trial stimulus synchrony. Previously we found (Eggermont 1994) that for noise-burst stimulation, the shift predictor accounted on average for less than one-third of the neural synchrony for dual electrode pairs but for ~3/4 in case of single electrode pairs. The result after shift-predictor correction was that the neural correlation under stimulus conditions was apparently lower than under spontaneous conditions (3 vs. 5%). This was attributed to a violation of the assumption of additivity of spontaneous spikes and stimulus induced spikes that forms the basis for the correction procedure. In the example shown in Fig. 3, the correlation and prediction were for a dual electrode pair, and our finding that the predictor contributed only one-quarter of the total synchrony is in the same range as our previous average result of one-third. This suggests, however, that the synchrony is largely the result of within-trial covariation in firing rate of the units, rather than the effect of stimulus-onset locked changes in neural connectivity that are predicted from the between-trial covariation in firing rate (Eggermont and Smith 1995b, 1996a).

The peak coincidence measure that we used throughout the analysis is not normalized for firing rate so neuron pairs with high firing rates will show in general a high number of peak coincident firings. This makes it difficult to distinguish clearly how much information is coded in firing rate and separately in the correlation of spike firings (i.e., calculated as a correlation coefficient) (Eggermont 1994). But even with this confound, significant differences in tuning curve bandwidth are obtained. Furthermore, the only thing that matters to allow this as a neural code is what the animal does with the spikes. Are they integrated during a 100-ms time window (as in an estimate of firing rate), is a correction for this overall firing rate performed by the animal, or are they used in mechanisms sensitive to the number of coincident firings that arrive at a cortical cell? As Abeles (1991) has argued, synchronously arriving spike activity is ~10-fold more effective than asynchronous arriving activity. This, clearly, may contribute to their relative arrival times.

**Is population vector coding a model for azimuth representation in auditory cortex?**

Although the population vector is only one of the possible read-out mechanisms of cortical activity, it is useful for estimating the amount of information available in “spike count” and “spike synchrony” representations of azimuth. Several requirements for the use of such a read-out mechanism exist. One of the more important ones is that all neurons in the area are contributing to the prediction. This may not be justified for AI because the representation of azimuth in AI could be patchy because of the requirement to map a large number of stimulus features onto the same two-dimensional surface (Schreiner 1995). This would imply that some neurons might not contribute to a population vector. In this respect, a sensory cortical area may be different from motor areas for which the technique obviously works (Georgopoulos et al. 1986, 1988). One of the other requirements for the validity of the population-vector model is that the neurons in the population are firing independently. As we have shown, the firings of simultaneously recorded neurons are synchronous under the stimulus conditions in our experiments. A previous study using noise-burst stimulation (Eggermont 1994) demonstrated that the amount of neural synchrony, expressed as a peak cross-correlation coefficient, in cat AI was on average 14.2% for single electrode pairs and 3.7% for dual electrode pairs. For single electrode pairs, this was in the same range as the value of 12% found in visual area MT (Zohary et al. 1994). Thus sensory cortical neurons are not firing independently, and this must reduce the effectiveness of a population code for representing sound azimuth. Under stimulus conditions and for an average correlation coefficient of 10%, the beneficial effect of pooling neuronal data on the signal-to-noise ratio, and thus on performance, is lost for population sizes exceeding 50–100 neurons (Zohary et al. 1994). Because most of the 102 neurons used in constructing the population vector were recorded on separate electrodes for which the peak cross-correlation coefficient is at most 5%, the upper limit of the population size has likely not been reached for our data and a better performance might be expected with larger samples.

Theoretically, a uniform distribution of preferred azimuths is a second requirement to assure that the population vector will point in the same direction as the sound azimuth (Georgopoulos et al. 1988; Salinas and Abbott 1994). Such a uniform distribution was obviously not present in our individual unit data but was created artificially by averaging all
units with the same preferred azimuth at 25 dB above threshold. A comparison of the predictions based on the average group data and individual neuron data (unpublished results) showed that the only differences occurred for frontal azimuth and 22.5° into the contralateral field, which were both predicted as ipsilateral.

For good prediction performance, the population vector also requires radially symmetric tuning curves, such as Gaussian, cosine, or cosine squared. The procedure’s main sensitivity is for the tuning curve width and the level of the background activity, i.e., that part of the firing rate that is not azimuth dependent (Georgopoulos et al. 1988; Seung and Sompolinsky 1993; Tanaka and Nakayama 1995). The root-mean-square (RMS) error in the prediction for neurons with optimal bandwidth, which is 90° for a cosine tuning curve, appears to depend both on the signal-to-noise ratio (peak response at best azimuth vs. lowest response level) and on the population size (Tanaka and Nakayama 1995). For our sample size of ~100 neurons and signal-to-noise ratios between 0.5 and 0.8, this translates theoretically into RMS errors between 30 and 80°. Our actual average RMS errors for the group were 32° for synchronous count at high intensities and 34° and 45° for peak firing rate at high and low intensities respectively. For individual units, the RMS error for the prediction over all intensities combined was 38°. So our findings are in the same range as theoretically predicted. This suggests that the weak directionality in the firings of AI neurons is one of the major reasons for the relatively poor performance of the population vector.

We stimulated in the frontal half field only, and this also has an effect on the accuracy at which far contralateral and far ipsilateral azimuths can be predicted with the population vector method. Assuming a symmetry around the −90 and +90° azimuths, we assigned values for azimuths at −112.5 and −135° based on the data for −67.5 and −45° and also for the ipsilateral equivalents and calculated the change in performance. The result was a reduction in the RMS error to only 20°, suggesting that a 360° speaker array, as for instance used by Middlebrooks et al. (1994) or a complete surround virtual sound field (Brugge et al. 1996) may result in improved prediction. With the inclusion of more “simulated” speaker data at positions −157.5 and +157.5°, the performance deteriorated below that for the original −90 to 90° range. Clearly the symmetry assumption does not apply for those locations as was also suggested by a SU example shown in Middlebrooks et al. (1994).

We assigned best azimuths only from a list of nine sound azimuths instead of using the actual values resulting from cosine curve fits. The largest differences occurred for far contralateral and ipsilateral azimuths, and this could have pulled the predictions more toward −90° and +90°. Overall we believe this effect to be small compared with the effect of some of the other limitations in the prediction.

The choice of the tuning curve’s best azimuth at 25 dB above threshold or using the azimuth for which the best response occurred across all intensities had little impact on the predictions. Intensity, however, has quite an impact, and our study suggests that at low intensities (<40 dB SPL), where the pinna effects dominate and nearly all units show a strong response limited to the contralateral field, the amount of information in the neural responses is too limited for useful predictions. For intensity levels above 40 dB SPL, a sensitive discrimination around the frontal azimuth is possible, as earlier suggested on the basis of interaural level disparities in SU firing rate (Phillips and Brugge 1985). This would suggest a two-step discrimination of the sound azimuth: first a coarse orientation toward contra or ipsi and second, on passing to the frontal plane, a fine tuning toward the sound source. Behavioral data (May and Huang 1996) suggest that for short noise bursts, cats showed the most accurate head-orientation responses in the frontal field. Presenting another noise burst after 3–5 s improved total response accuracy by allowing correction movements. This is in qualitative agreement with the predictions at stimulus levels >40 dB SPL. At low-intensity levels, head scanning (or pinna movement) would be needed for precise localization based on information in AI.

Our predictions are sometimes slightly outside the range from −90 to +90°, and this is a consequence of using only the azimuth dependent part of the firing rate arrived at by subtracting the mean firing rate (or synchronous rate) across all azimuth conditions. So for a −90° azimuth, the contribution of, say, the 67.5° best azimuth units is generally negative and its contribution is thus in the −112.5° direction. This happens for other ipsilaterally best azimuths as well, and the general result is that the “vote” for the predicted azimuth will be for angles larger than −90°. The same applies for far ipsilateral azimuth predictions. The effects cancel for azimuths around the midline.

Implications for the role of AI in azimuth coding

Given that, at intensities >40 dB SPL, the population vector method suggests that insufficient information is available in the two simple measures of population activity of AI, firing rate, and peak firing synchrony, to allow reasonably accurate estimations of sound azimuth, one wonders why there is a contralateral deficit after lesioning one AI (Jenkins and Merzenich 1984; Kavanagh and Kelly 1987). The ipsilateral AI is likely to have a mirror representation of the contralateral AI, implying that for sound localization only contralaterally tuned units matter (Phillips and Brugge 1985). These are the units that have the shortest latencies, represent the core of the spatial receptive field, and persist at relatively low-intensity levels (Brugge et al. 1996; Eggermont 1998). Could it be that lesioning AI in one hemisphere deprives the animal only from making decisions along the midline, i.e., of not making orienting head or pinna movements? A reduction in the number of corrective head movements has been observed only after bilateral lesioning of the entire auditory cortex (Beitel and Kaas 1993). In the same study, it was shown that unilateral lesions of auditory cortex in cat did not abolish correct orienting responses to the left or right hemifield, and it was argued that separate pathways exist for acoustical orientation (a reflexive task) and sound localization requiring an association of sound with a localization in space (a cognitive task). Heffner and Masterton’s (1975) findings in macaque monkeys with bilateral cortical lesions also suggest that the effects are not purely sensory but may involve disruption of sensorimotor integration. Thus either sound orienting is based on subcortical mechanisms or the patency of one AI is sufficient for its execution.
whereas two AIs, or pathways passing through them, are required for the more demanding task of sound localization.

Firing rate and firing synchrony appear to carry the same information in the context of the population vector coding model, so that predictions of sound azimuth based on these data are indistinguishable at intensities >40 dB SPL. At low intensities, peak firing rate and synchronization between spikes and LFPs perform very poor. There are two potential reasons for this. One is the smaller size of the LFPs at these intensities resulting in a increased jitter in the triggers and frequent absence of triggers. The second reason is that LFPs at these low levels are only obtained for contralateral azimuths and as a result synchronization with spikes produced by frontal and ipsilateral azimuths is absent. As a consequence, predictions on basis of neural synchrony at low levels perform in a binary way: absence or presence of synchrony corresponding to contralateral versus ipsilateral predictions. Thus population-vector decoding suggests that especially at higher intensities there is a crude functional representation of azimuth in AI that contains sufficient information for head orientation to the contralateral or ipsilateral hemifield and provides good spatial resolution for azimuths within 22.5° from the midline.

Middlebrooks et al. (1994) have shown that a single-layer linear neural network in a supervised learning paradigm that makes use of the probability of firing in 40 successive 1-ms bins after stimulus onset and calculates a weighted sum of the spike probability could be trained to on average perform a better prediction than the population vector method used in this study. The population-vector method only uses one number per response such as peak count, overall spike count, synchronized count, etc. and does not include temporal information contained in the burst of activity after the stimulus onset. As can be seen from the pooled PSTHs in Fig. 10, especially at intermediate intensities (35 and 45 dB SPL), there is a clear distinction in the average distribution of spikes across the response window for different azimuths. This disappears for higher intensities. If the neural network procedure that was successful in decoding the activity in AES also could apply to AI, then the distribution of relative spike times would be expected to provide the extra information, especially for the far contralateral and far ipsilateral azimuths, allowing an on average better prediction. However, response patterns for the three contralateral azimuths from −45 to −90° are generally very similar as shown also in the dot displays and PST histograms (Fig. 1). The pooled data in Fig. 10 also suggest that a prediction procedure based on temporal aspects within a response window might perform best at intermediate intensities (35–45 dB SPL) where the temporal variability in the responses is largest.

G. Smith provided software support and made suggestions throughout the experiment. J. Schnupp commented on a previous version of the manuscript. D. Bowman, K. Ochi, and M. Kenmochi assisted with the data collection.

This investigation was supported by grants from the Alberta Heritage Foundation for Medical Research and the Natural Sciences and Engineering Research Council of Canada.

Address for reprint requests: J. J. Eggermont, Dept. of Psychology, The University of Calgary, 2500 University Dr., N.W., Calgary, Alberta T2N 1N4, Canada.

Received 16 October 1997; accepted in final form 8 July 1998.

REFERENCES


