Codes for Sound-Source Location in Nontonotopic Auditory Cortex

JOHN C. MIDDLEBROOKS,[1,3] LI XU,[1,3] ANN CLOCK EDDINS,[1] AND DAVID M. GREEN[2]
[1]*Department of Neuroscience and* [2]*Department of Psychology, University of Florida, Gainesville, Florida 32610;*
and [3]*Kresge Hearing Research Institute, University of Michigan, Ann Arbor, Michigan 48109-0506*

Middlebrooks, John C., Li Xu, Ann Clock Eddins, and David M. Green. Codes for sound-source location in nontonotopic auditor cortex. *J. Neurophysiol.* 80: 863–881, 1998. We evaluated two hypothetical codes for sound-source location in the auditory cortex. The topographical code assumed that single neurons are selective for particular locations and that sound-source locations are coded by the cortical location of small populations of maximally activated neurons. The distributed code assumed that the responses of individual neurons can carry information about locations throughout 360° of azimuth and that accurate sound localization derives from information that is distributed across large populations of such panoramic neurons. We recorded from single units in the anterior ectosylvian sulcus area (area AES) and in area A2 of α-chloralose–anesthetized cats. Results obtained in the two areas were essentially equivalent. Noise bursts were presented from loudspeakers spaced in 20° intervals of azimuth throughout 360° of the horizontal plane. Spike counts of the majority of units were modulated >50% by changes in sound-source azimuth. Nevertheless, sound-source locations that produced greater than half-maximal spike counts often spanned >180° of azimuth. The spatial selectivity of units tended to broaden and, often, to shift in azimuth as sound pressure levels (SPLs) were increased to a moderate level. We sometimes saw systematic changes in spatial tuning along segments of electrode tracks as long as 1.5 mm but such progressions were not evident at higher sound levels. Moderate-level sounds presented anywhere in the contralateral hemifield produced greater than half-maximal activation of nearly all units. These results are not consistent with the hypothesis of a topographic code. We used an artificial-neural–network algorithm to recognize spike patterns and, thereby, infer the locations of sound sources. Network input consisted of spike density functions formed by averages of responses to eight stimulus repetitions. Information carried in the responses of single units permitted reasonable estimates of sound-source locations throughout 360° of azimuth. The most accurate units exhibited median errors in localization of <25°, meaning that the network output fell within 25° of the correct location on half of the trials. Spike patterns tended to vary with stimulus SPL, but level-invariant features of patterns permitted estimates of locations of sound sources that varied through 20-dB ranges. Sound localization based on spike patterns that preserved details of spike timing consistently was more accurate than localization based on spike counts alone. These results support the hypothesis that sound-source locations are represented by a distributed code and that individual neurons are, in effect, panoramic localizers.

INTRODUCTION

An intact auditory cortex is essential for normal localization of sounds. Cortical pathology in humans results in deficits in sound localization ability (Greene 1929; Klingon and Bontecou 1966; Sanchez-Longo and Forster 1958; Wortis and Pfeiffer 1948). Similarly, unilateral ablations of the auditory cortex in animals produce behavioral deficits in localization of sound sources presented on the side contralateral to the lesion (Jenkins and Masterton 1982). Neurophysiological studies of the optic tectum in the barn owl and the superior colliculus in mammals show that single neurons are selective for sound-source location (barn owl: Knudsen 1982; guinea pig: Palmer and King 1982; cat: Middlebrooks and Knudsen 1984; monkey: Jay and Sparks 1984; ferret: King and Hutchings 1987). In those midbrain structures, neurons' preferred sound-source locations vary systematically according to the locations of the neurons within the structure. These two sets of observations, that auditory cortex is essential for localization and that the nervous system is capable of forming a map of auditory space (at least in the midbrain), strongly suggest the hypothesis that sound-source locations are represented topographically in the mammalian auditory cortex.

Surprisingly, efforts in several laboratories to find maps of auditory space in the cortex have produced disappointing results. Previous studies have examined cortical areas A1 in cat and monkey (cat: Brugge et al. 1994, 1996; Imig et al. 1990; Middlebrooks and Pettigrew 1981; Rajan et al. 1990b; monkey: Ahissar et al. 1992) and, to a lesser degree, the cat's anterior ectosylvian area (area AES) and anterior auditory area (Korte and Rauschecker 1993). Those studies have shown that a subset of auditory cortical neurons can exhibit modulation of spike counts by changes in sound-source azimuth. Many of the ''high directional'' neurons, however, respond strongly to sounds presented across areas as large as half the sound field. Moreover, studies consistently have shown that the spatial tuning of most neurons broadens considerably as the stimulus intensity is increased to more than ~20 dB above the neuron's threshold. Neurons recorded successively along an electrode track sometimes show systematic shifts in spatial tuning as a function of unit location. Nevertheless, such sequences have not been shown to extend more than a fraction of the width of an auditory cortical area without interruption by neurons that show very different spatial sensitivity (Clarey et al. 1994; Imig et al. 1990; Middlebrooks and Pettigrew 1981; Rajan et al.1990a). If sound-source locations are represented in the auditory cortical areas that have been studied so far, the form of the representation must be very different from the auditory space maps that have been demonstrated in the optic tectum and superior colliculus.

In this study, we explored spatial representation in area AES and area A2 of the cat's auditory cortex. Those areas are ''nontonotopic'' in the sense that they do not show an obvious topographic representation of sound frequency. We chose those areas for this study because neurons there are known to respond well to sounds that have broad bandwidths, and broadband sounds are localized more accurately than tones (e.g., Stevens and Newman 1936). Also, area

AES is the only auditory cortical area in the cat that has been shown to project strongly to the superior colliculus (Meredith and Clemo 1989), which contains an auditory space map. We evaluated two hypothetical codes for sound-source location: a *topographical* code and a *distributed* code. The topographical code hypothesis assumed that each neuron is selective for a particular sound-source location, that the preferred locations of neurons vary according to cortical location, and that the location of a sound source is coded by the location in the cortex of a small population of maximally active neurons. We tested that hypothesis by plotting conventional spike-count-versus-azimuth profiles of units and by searching for systematic shifts in spatial tuning as a function of cortical location. The results for areas AES and A2 were qualitatively quite similar and, in turn, were similar to those described in published studies of area A1. Many neurons showed clear spatial tuning at low sound pressure levels (SPLs), but the topography was fragmentary and tended to degrade further at moderate SPLs. Our results are not consistent with the topographical code hypothesis.

The distributed code hypothesis assumed that the activity of individual neurons can carry information about broad ranges of location and that accurate sound localization is derived from information that is distributed across large populations of neurons. We tested that hypothesis by attempting to recognize the firing patterns of neurons that resulted from each source location and, thereby, read the locations of sound sources from the neural firing patterns. Most previous studies have represented the responses of units only by the magnitudes of responses (i.e., spike counts or rates), but we employed artificial neural networks to recognize complete spike patterns, which included the timing of spikes as well as spike counts. We found that, for the majority of units, spike times carried substantial stimulus-related information beyond that carried by spike counts alone. Our results show that the firing pattern of a single neuron could code the location of a sound source, with varying degrees of accuracy, throughout 360° of azimuth. Some features of spike patterns changed with changes in stimulus pressure level, but level-invariant features of spike patterns permitted localization of sounds that vary in sound level. These results support the hypothesis that sound-source locations are represented in the auditory cortex by a distributed code.

METHODS

Experimental apparatus and stimulus generation

The series of experiments was begun at the University of Florida and concluded at the University of Michigan. The sound chamber and facilities for stimulus generation and data recording were essentially equivalent at the two institutions. Experiments were controlled with an Intel-based personal computer. Acoustic stimuli were synthesized digitally, using equipment from Tucker-Davis Technologies (TDT). The sample rate for audio output was 100 kHz, with 16-bit resolution. Experiments were conducted in a sound-attenuating chamber that was lined with acoustical foam (Illbruck) to suppress reflections of sounds at frequencies >500 Hz. Sounds were presented from multiple loudspeakers, one loudspeaker at a time, from a distance of 1.2 m from the animal; the speakers were Pioneer model TS-879 two-way coaxials. A circular hoop held 18 loudspeakers in the horizontal plane with angular separation of 20°. A second hoop held 14 loudspeakers in the vertical midline plane with angular separation of 20° from 60°

below the frontal horizon, up and over the top, to 20° below the rear horizon. A computer-controlled multiplexer (TDT model PM1) permitted any one loudspeaker to be activated at any time. The loudspeakers were calibrated by presenting maximum-length sequences (Golay codes) (Zhou et al. 1992) and recording the responses with a precision microphone (Bruel and Kjaer, model 4133) placed in the center of the chamber in the absence of the cat. Loudspeaker responses were equalized individually so that the root-mean-squared variation in sound level, computed in 6.1-Hz steps from 1,000 to 30,000 Hz, was <1.0 dB.

Noise bursts were used to measure spatial sensitivity of units. An independent Gaussian noise sample was used for each stimulus presentation, rather than repeating a constant ''frozen'' noise sample. This was necessary to avoid entrainment of neural responses to the envelope of frozen noise, which might have produced erroneous time structure in unit firing patterns. The spectra of the Gaussian noise bursts were band-passed between 1 and 30 kHz with abrupt cutoffs. Noise-burst durations were 80–300 ms, except when stated otherwise, and had abrupt onsets and offsets. Tone bursts were used to measure the frequency sensitivity of units. Tone levels were calibrated for the sound field in the absence of the cat (i.e., not at the cat's tympanic membrane), so measurements of unit frequency sensitivity were influenced by the acoustical properties of the external ears. Tone bursts were 80–100 ms in duration, ramped on and off with 5-ms rise/fall times. Noise and tone bursts were presented once every 800 or 1,000 ms.

Animal preparation and unit recording

This report presents data from purpose-bred adult cats of both sexes. Data were obtained from 169 units in area AES in 14 cats and from 62 units in area A2 in 5 additional cats. Partial data from 55 of the AES units have appeared previously (Middlebrooks et al. 1994), but those data have been entirely reanalyzed for this report. Each cat was anesthetized for surgery with isoflurane in 70% N$_2$O and 30% O$_2$. The concentration of isoflurane was adjusted so that limb withdrawal reflexes were abolished. Cats were transferred to α-chloralose anesthesia for unit recording. The induction time for α-chloralose anesthesia was ~3 h, so intravenous injections of a solution of α-chloralose (25 mg/ml in propylene glycol) were begun immediately after induction of the gas anesthesia. A typical loading dose of α-chloralose was 125–150 mg. Typically, ~2 h passed between the end of isoflurane administration and the beginning of unit recording, so we presume that the major anesthetic effect during data collection was due to the α-chloralose. During unit recording, supplemental injections of α-chloralose were given whenever a strong pinch of the forepaw resulted in a prolonged elevation of heart rate. An esophageal stethoscope fitted with a thermometer was used to monitor heart rate and core temperature. A warm-water heating pad was used to maintain temperature at 38°C. Ringer solution was given intravenously at a rate of ~10 ml/h to maintain hydration.

All recordings were made from the right cortical hemisphere. A midline scalp incision was made and the temporalis muscle was retracted on the right side. Portions of scalp and temporalis muscle were removed to make room for the recording chamber. A stainless-steel fixture was attached to the skull with screws and dental cement. A skull opening was made to reveal the middle ectosylvian gyrus and anterior ectosylvian sulcus and a plastic chamber was cemented around the opening to contain a pool of silicone oil. The scalp was sutured closed around the plastic chamber. The animal was transferred to the center of a sound-attenuating chamber, with its interaural axis centered in the sound chamber, 1.3 m above the floor. The animal's body was supported in the heating pad in a sling, and its head was supported from behind by a bar attached to the skull fixture. Thin wire supports were used to push the external ears into a forward position (Middlebrooks and Knudsen

1987). The position of the ears was constant throughout each experiment.

Unit activity was recorded with parylene-insulated tungsten microelectrodes (Frederick Haer); nominal impedances were ~4 MΩ. Activity was amplified, and spikes were discriminated on-line with an amplitude and time discriminator (TDT model SD1). Whenever possible, the discriminator was adjusted to isolate single units, but in the worst cases, the discriminator probably accepted two or more indiscriminable units. We presume that contamination of single-unit recording by additional units could only increase the apparent breadth of spatial tuning and could only decrease the spatial specificity of spike patterns. For that reason, we regard our results to be conservative estimates of the accuracy of spatial coding by single units. Spike times were digitized and stored with 100-μs resolution. Custom graphics software provided on-line display in the form of raster plots, poststimulus time histograms, and bar plots of spike counts versus various stimulus parameters. Study of each unit took ~2 h.

Recordings from area AES were made from electrode tracks that passed down the posterior bank of the anterior ectosylvian sulcus. Most tracks began near the dorsomedial tip of that sulcus and yielded units along ~4 mm of the track. Recordings from area A2 were made from penetrations that passed oblique to the cortical surface near the crest of the middle ectosylvian gyrus, ventral to area A1. Search stimuli consisted of broadband noise bursts, presented in the region of 0° to contralateral 40° azimuth. Area AES was distinguished from the anterior auditory field, and area A2 was distinguished from cortical area A1, by the absence of tonotopic organization and by 40-dB bandwidths that were one or more octaves. Frequency tuning is considered in greater detail in the companion paper (Xu et al. 1998). Electrode tracks were marked with electrolytic lesions at the ends of most tracks and at one or more depths as the electrode was withdrawn. Experiments typically lasted 30–60 h and yielded 9–18 units along one or two electrode tracks in area AES or ~14 units along one to three tracks in area A2.

At the end of most experiments, the animal was killed with a lethal dose of pentobarbital sodium or potassium chloride (intravenous) and then was perfused transcardially with buffered aldehydes. The brain was sectioned and stained with cresyl violet to localize electrode tracks.

Experimental procedure

Study of each unit began by identifying a sound-source azimuth at which the unit responded reliably, typically 0° or contralateral 40°, then measuring responses to noise bursts at a range of SPLs in 5-dB steps. The unit's threshold was estimated to the nearest 5 dB by inspection of poststimulus time histograms and bar plots of spike counts versus SPLs. Then the unit's spatial sensitivity was measured using a stimulus set that typically consisted of noise bursts presented from 18 azimuths in the horizontal plane (−180° to 160° in steps of 20°) at 2 or 5 SPLs ranging from 20 to 40 dB above the unit's threshold. In some instances, stimuli in the horizontal plane were interleaved with stimuli at various elevations in the vertical midline plane. Elevation sensitivity of units is considered in the companion paper (Xu et al. 1998). Stimuli were presented in pseudorandom order such that all locations were tested at all SPLs once before repeating all stimuli again in a different random order. Each combination of location and SPL was tested ≥40 times. Frequency sensitivity was measured with a sound source fixed at a location at which a noise source produced a strong response, usually 0° or contralateral 40° azimuth. Tone frequencies were varied in one-third–octave steps from 3.75 to 30 kHz.

Data analysis

Spike times were stored with 100-μs resolution as latencies relative to the estimated time of arrival of sound at the center of the sound chamber, assuming an acoustic travel time of 4 ms. Spike patterns were expressed with 1-ms resolution by convolving spike times with a Gaussian impulse ($\sigma = 1$ ms), then resampling at 1 kHz. Convolution with the Gaussian impulse served to low-pass filter the spike patterns below 137 Hz, thereby attenuating aliased high frequencies, and served to smooth the otherwise-sparse spike density functions that were used as input to the artificial neural network. For the purpose of testing the artificial neural network recognition of spike patterns, we divided responses into training and test sets. From the set of all responses to a particular stimulus, the odd-numbered trials were assigned to the training set and the even-numbered trials were assigned to the test set. The separation of training and test sets provided a cross-validation of the pattern-recognition scheme. A bootstrap averaging procedure (Efron and Tibshirani 1991) was used to form average spike density functions within the test set or within the training set. Given a training set (or a test set) of 20 responses to each stimulus condition, we formed each density function by repeatedly drawing eight samples with replacement from the set of spike patterns elicited by stimuli of particular location and SPL. Because we sampled with replacement, each bootstrap average could contain no, one, two, or more instances of each spike pattern. The bootstrap procedure was used to estimate the variability in the averages, given the limited number of responses measured at each location. For each unit, we formed 20 bootstrapped training patterns and 100 bootstrapped test patterns for each stimulus condition.

Artificial neural networks were constructed with the MATLAB Neural Network Toolbox (The Mathworks, Natick, MA). Supervised training of the networks used the back-propagation algorithm (Rumelhart et al. 1986). The training procedure incorporated Nguyen-Widrow initial conditions, momentum, and an adaptive learning rate (Demuth and Beale 1995). During training, the network was presented only with spike patterns in the training set. Overtraining with the training set would have led to increases in the error in recognition of the test set. We avoided overtraining by periodically testing accuracy of recognition of the test set. Training was halted when errors in recognition of the test set began to increase. The back-propagation algorithm is a gradient-descent procedure that begins with randomized weights and biases. Therefore repeated training of networks using a given set of data produced slightly varying outputs. For that reason, we repeated the network training three times with the training set of responses from each neuron, then recorded the output of the network that produced the smallest median error. The network architecture was similar to the feed-forward architecture that was preferred in a comparison of network architectures for study of the visual cortex (Kjaer et al. 1994). The main difference was that our network produced a scalar estimate of the stimulus, whereas the one described by the Kjaer group produced an output that was quantized to particular stimuli and then was used to compute transmitted information. In our study, the central tendency of multiple estimates of stimulus locations was represented by the *mean direction* (Fisher et al. 1987). The mean direction was computed by treating each estimated location as a unit vector, forming the vector sum, then finding the direction of the resultant vector.

An analysis of variance (ANOVA) procedure (Hays 1981) was used as a method independent of artificial-neural–network analysis to quantify the degree by which spike patterns were modulated by sound-source azimuth. Variance across spike patterns was found by computing the variance in spike density in each 1-ms time bin, then summing the variances across all time bins. Variance was computed before and after sorting spike patterns according to azimuth. The result of the procedure was the percent variance accounted for by azimuth.

RESULTS

We recorded from units in cortical area AES on the posterior bank of the anterior ectosylvian sulcus, and in cortical

area A2 near the crest of the middle ectosylvian gyrus, ventral to area A1. We begin this report by presenting conventional measures of azimuth sensitivity of spike counts. Next we examine azimuth coding by unit spike patterns. That analysis makes use of artificial neural networks for pattern recognition. We consider the issue of panoramic coding of sound-source location by single neurons. Finally, we compare azimuth coding by spike patterns with azimuth coding by spike counts alone.

Azimuth sensitivity of spike counts

Most units showed some degree of spatial tuning in that the number of spikes elicited by a noise burst varied as a function of the sound-source location. Figure 1 shows polar plots of mean spike counts versus azimuth for two units in area AES (Fig. 1, *A* and *B*) and two in area A2 (Fig. 1, *C* and *D*); *left* and *right columns* show responses at sound pressures of 20 and 40 dB, respectively, above threshold. The unit shown in Fig. 1*A* was typical of many units in areas AES and A2 in that it responded well to sound sources throughout the hemifield contralateral to the recording site. The unit in Fig. 1*B* showed two or three peak responses at the lower sound level that resolved to a single peak at the higher level. The unit in Fig. 1*C*, had it been tested only with low-level sounds in the frontal hemifield, would have appeared to be tuned to the frontal midline, but when tested with sound sources throughout 360° of azimuth it showed strong responses to sounds behind the cat. The unit shown in Fig. 1*D* is an example of a unit that showed fairly broad spatial sensitivity.

We quantified the azimuth sensitivity of unit spike counts by computing the depth of modulation of spike counts by azimuth (Fig. 2). For noise bursts that were 20 dB above units' thresholds, 94% of units in area AES (145 of 154 units) and 77% of units in area A2 (48 of 62 units) showed >50% modulation of their spike counts across an azimuth range of 360°. Modulation depths decreased at sound levels 40 dB above threshold. The distribution of modulation depths was qualitatively similar between areas AES and A2, although the depth of modulation was significantly less in area A2 at both sound levels ($P < 0.01$, Mann-Whitney U test). Despite the strong modulation of spike counts by sound-source azimuth, spatial tuning generally was quite broad. This is represented in Fig. 3 by the ranges of azimuth across which units responded with >50% of their maximum response rates. When sound levels were 20 dB above threshold, 58% of units in area AES (89/154) and 81% of units in area A2 (50/62) responded with greater than half-maximal spike counts to sound sources throughout ≥180° of azimuth. Consistent with the changes in modulation depth, the widths of half-maximal response areas increased as sound levels were increased from 20 to 40 dB above threshold.

Given the broad half-maximal response areas of most units, it seemed likely that a single sound source would activate a large percentage of the unit population. To estimate that percentage, we plotted as a function of sound-source azimuth the percentages of our unit samples that were activated above spike-count criteria of 25, 50, and 75% of maximum (Fig. 4). Those plots demonstrate a strong contralateral bias in the tuning of most neurons,
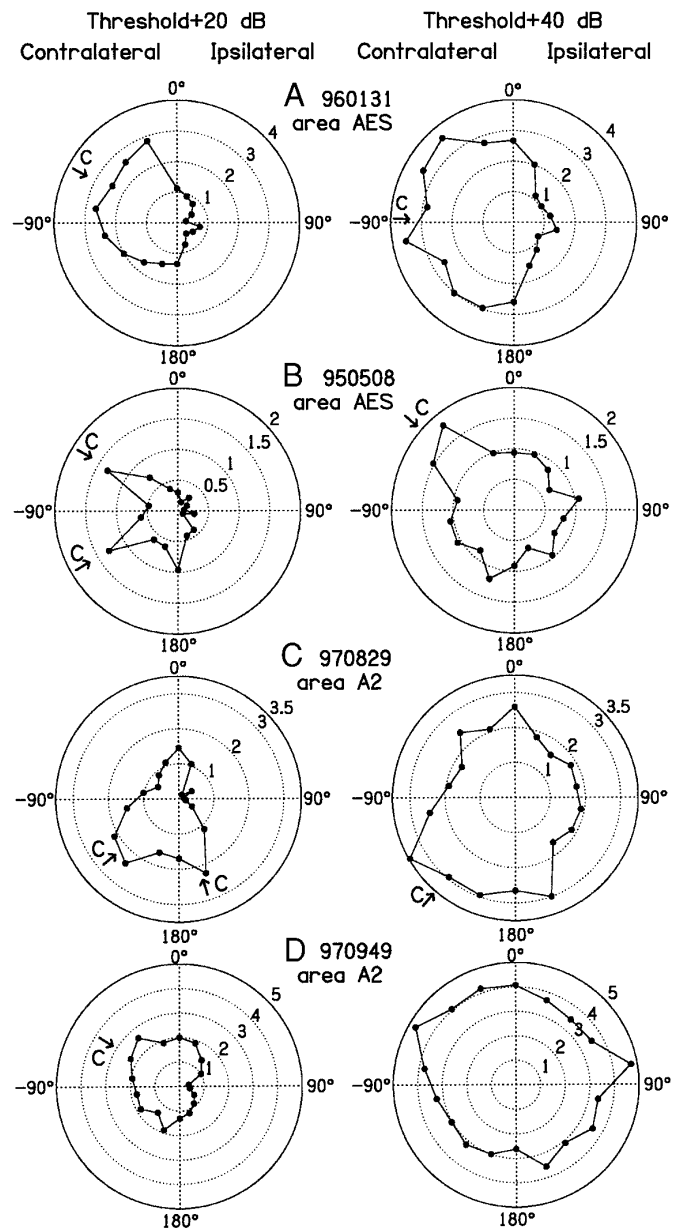

FIG. 1. Spike-count-versus-azimuth profiles. Each *horizontal row* of 2 panels represents azimuth profiles for 1 unit at 2 sound levels. *Left* and *right*: profiles at 20 and 40 dB, respectively, above units' thresholds. In these polar plots, the angular dimension gives the location in azimuth of the sound source in the horizontal plane, with 0° straight in front of the cat and negative values to the cat's left, contralateral to the recording site in the right cortical hemisphere. Radial dimension gives the mean spike count, expressed as spikes per stimulus presentation. Arrows labeled ''C'' indicate the best-azimuth centroids as defined in the text.

something that has been observed in most studies of the auditory pathway at or above the level of the inferior colliculus. More interesting, the plot suggests that, across a broad range of contralateral azimuths, neurons are limited in the dynamic range of spike counts with which they can discriminate among azimuths. That is, a noise burst 40 dB above each unit's threshold activated most of the neurons within our sample to at least half of their maximum spike counts when the source was located anywhere on the side contralateral to the recording site. Similarly, contralateral stimuli at that pressure level produced >75%
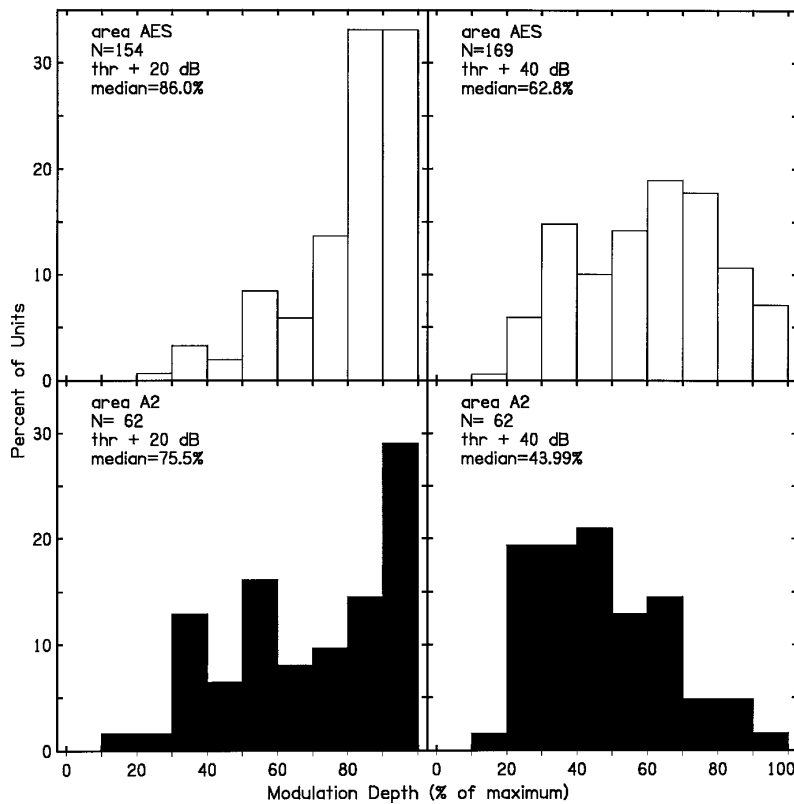
FIG. 2. Modulation of spike counts by sound-source azimuth. Spike-count modulation is given as the maximum percentage reduction of spike count as a constant-level sound source was varied in location through 360° of azimuth [i.e., $100 \times (1 - \text{min/max})$]. *Top* and *bottom*: sampled unit population in areas anterior ectosylvian sulcus (AES) and A2, respectively. *Left* and *right*: measurements made with stimuli 20 and 40 dB, respectively, above units' thresholds.

activation of more than half of our sample. Most of the units in our sample had thresholds in the range of $-5-25$ dB SPL, so a level of 40 dB above threshold is only a moderate sound level. These data suggest that a model of azimuth coding based only on spike counts would require that the majority of units code locations throughout most of contralateral azimuth with only ~25% of their dynamic ranges.
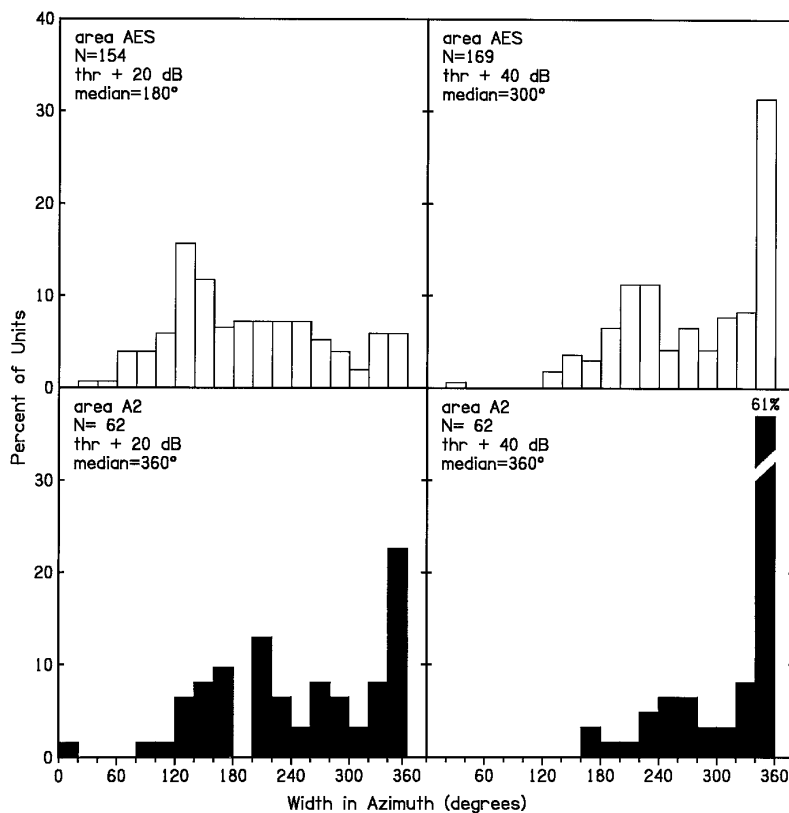

FIG. 3. Width of azimuth tuning. Azimuth tuning was tested with constant-level sound sources varied in 20° increments of azimuth. Width of azimuth tuning is given by 20° times the total number of tested azimuths at which at least a half-maximal spike count elicited. *Top* and *bottom*: areas AES and A2, respectively. *Left* and *right*: 20 and 40 dB, respectively, above units' thresholds.
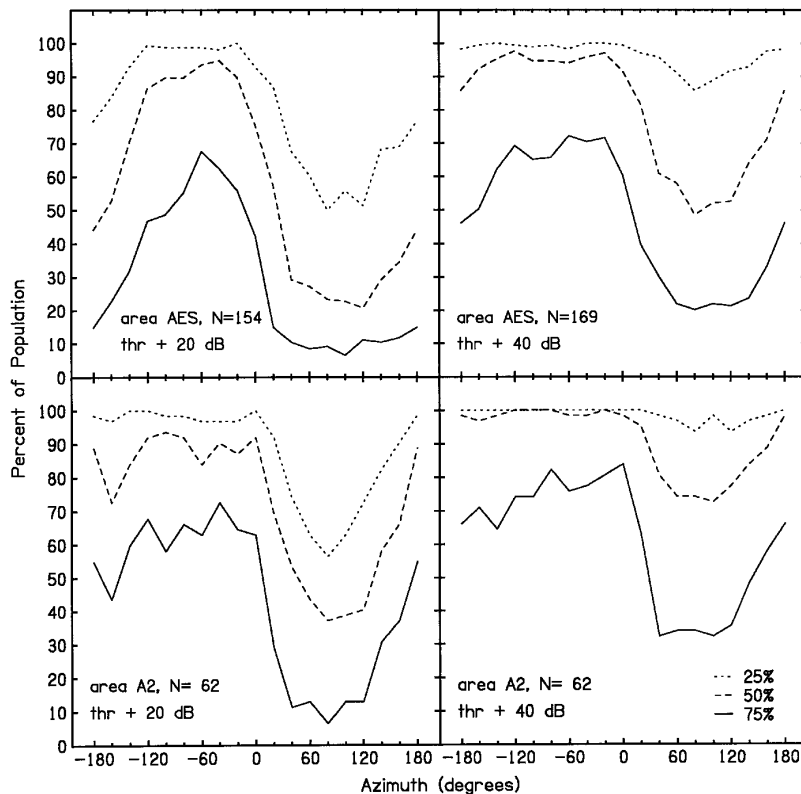
FIG. 4. Percentage of unit populations activated by sound sources at various azimuths. These plots represent normalized spike-count data from 154 or 169 units in area AES (*top*) and 62 units in area A2 (*bottom*). Three lines in each panel show the percentage of the sample population that was activated at or above 25, 50, and 75% of each unit's maximum spike count. Data are plotted as a function of sound-source azimuth. *Left* and *right*: 20 and 40 dB, respectively, above units' thresholds.

BEST-AZIMUTH CENTROIDS. We often encountered azimuth profiles in which the location of the single sound source that elicited the most spikes did not appear to represent accurately the directionality of the unit. For instance, Fig. 1*A* shows an azimuth profile in which the maximum response was obtained with a sound-source at contralateral 20° but in which the general azimuth preference of the unit seemed also to include locations further contralateral. Moreover, some units showed multiple peaks in their azimuth sensitivity, as in Fig. 1*B*. We attempted to represent the directional preference of units by the direction of the spike-count–weighted vector sum of all responses, but in the case of multiple peaks, the resultant vector often would point in the direction of a local minimum centered between peaks. Instead, we used the following procedure to compute one or two *best-azimuth centroids* from each azimuth profile. First, we selected units that showed ≥50% modulation of their spike rates. Second, we defined a *peak* as a set of one or more contiguous azimuths at which the mean responses were greater than a criterion of 75% of the maximum mean response. Third, we computed the vector sum of all of those azimuths plus the two subcriterion responses recorded on either side of the peak. In forming this vector sum, the vector representing each azimuth had a direction corresponding to the stimulus azimuth and a length corresponding to the spike count. Finally, the best-azimuth centroid was given by the direction of the resultant vector. Arrows labeled ''C'' identify centroids in Fig. 1. We favor centroids over conventional measures of ''best area centers'' (Imig et al. 1990; Knudsen 1982) or ''peak response azimuths'' (Rajan et al. 1990b) because the location of a centroid is weighted by all the measurements within a peak, not just by a single maximum. Also, our definition of centroids permitted us to deal with

multiple peaks in response profiles by computing multiple centroids. Figure 5 shows the distributions of centroids across our samples of units. Each unit is represented by the centroid of the tallest peak in the unit's azimuth profile. Centroids were distributed continuously throughout the contralateral half of space, and centroids of a few AES units were scattered on the ipsilateral side. When sound levels were 40 dB above units' thresholds, 31% of units in area AES and 61% of units in A2 showed modulation depths that were too shallow to permit computation of centroids (shown as NC in Fig. 5).

INFLUENCE OF SOUND PRESSURE LEVEL ON SPATIAL TUNING. Most units showed a substantial broadening of their spatial tuning as sound levels were increased from 20 to 40 dB above units' thresholds. This is shown by the differences between the *left* and *right columns* of Fig. 3. Spatial tuning widths at half-maximum spike counts were wider at 40 dB than at 20 dB above threshold for 88% of units in area AES and 77% of units in area A2. Not surprisingly, the locations of best-azimuth centroids of many units also shifted in azimuth with changes in sound level. Figure 6 compares the best-azimuth centroids of units measured at 20 and 40 dB above threshold. Data are presented only for the 102 units in area AES and 24 units in area A2 that showed ≥50% modulation of spike count at both sound levels, so the figure represents only the more azimuth-selective half of our sampled population. Many of the points lie near the diagonal line that indicates equal centroids at the two levels. Nevertheless, a third of the units in area AES (33%; 34/102) and 17% of units in area A2 (4/24) showed shifts in centroid location of >40°, which is more than two times the minimum loudspeaker separation that we used.
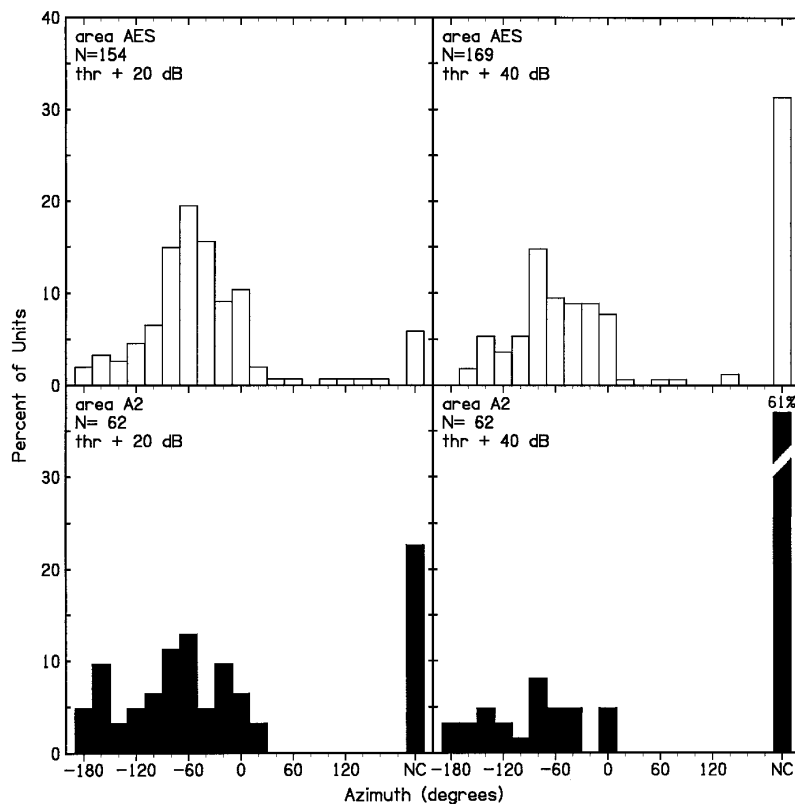
FIG. 5. Distribution of best-azimuth centroids. *Primary centroid* of each unit is the centroid of the tallest peak in its azimuth profile. NC (no centroid) represents units that showed <50% modulation of their spike counts by azimuth and, by our definition, had no measurable best-azimuth centroid. *Top* and *bottom*: areas AES and A2, respectively. *Left* and *right*: 20 and 40 dB, respectively, above units' thresholds.

MULTIPEAKED AZIMUTH PROFILES. A substantial number of azimuth profiles showed multiple peaks. For the purpose of quantification, we defined a "secondary peak" as two or more contiguous points in an azimuth profile that were >75% of a unit's maximum response. We required that a secondary peak be separated from a taller "primary peak" by at least one point at which the response was <50% of maximum or by at least two points at which the responses were <75% of maximum. At sound levels 40 dB above units' thresholds, 24% of units in area AES (40/169) and 15% of units in area A2 (9/62) showed secondary peaks in their azimuth profiles. We computed the centroids of secondary peaks as described above for primary peaks. Figure 7, *A* and *B,* shows the relation between the centroids of primary and secondary peaks. The dashed lines indicate pairs of azimuths that are mirror-symmetric with respect to the interaural axis. Such points are analogous to the "front/back confusions" that often are reported in behavioral studies (e.g., Stevens and Newman 1936; Wenzel et al. 1993). There are several instances in which the data points lie on or near the dashed line (i.e., primary and secondary centroids are mirror images), but there also are numerous examples of data points that deviate substantially from the line.

TOPOGRAPHY OF SPATIAL TUNING. We saw no consistent map of sound-source azimuth in the form of a systematic progression of azimuth centroids across cortical locations. Sometimes, however, we could observe local trends in centroid locations as a function of recording position. Our most productive electrode penetrations passed down the bank of the anterior ectosylvian sulcus, parallel to the cortical layers, and recorded mostly from units in the middle cortical layers. All penetrations, however, initially passed through superficial layers and most ended in deep layers. For that

reason, cortical depth is confounded somewhat with distance down the sulcal bank, and we cannot distinguish the contributions of those two factors in consideration of apparent trends in best-azimuth centroids. Figure 8 shows the locations of centroids as a function of recording depth for four electrode penetrations down the posterior bank of the anterior ectosylvian sulcus. These four penetrations are our best instances of orderly progressions of centroids along relatively long penetrations. The *left* and *right columns* of plots show centroids measured at sound levels of 20 and 40 dB, respectively, above thresholds (●, primary centroids of units; ○, secondary centroids when present; ×, units that showed <50% modulation of their spike counts by azimuth and, by our definition, had no centroids). One can see instances of smooth progressions of centroids. In some instances (e.g., Fig. 8*H*), a gap in a progression of primary centroids is filled in by one or more secondary centroids. We tested for systematic organization in centroid azimuth along electrode penetrations by computing the correlation of the azimuth centroid of each unit with the centroid of the unit recorded next along the same electrode penetrations; we tested only pairs of units that were separated by no more than 400 μm. We did this analysis in area AES, where penetrations tended to cross the cortical cell columns, and not in area A2, where penetrations often were roughly parallel to cortical columns. The correlations in area AES were $r = 0.59$ at 20 dB and $r = 0.51$ at 40 dB. Both correlations were highly significant ($P < 0.001$) although not particularly strong. A correlation of 0.59 means that knowledge of the best-azimuth centroid of a unit reduces by 35% the variance in the centroid of the next recorded unit. In the data shown in Fig. 8 for sound levels 20 dB above threshold, one can see a suggestion of a pattern of near-midline centroids at the
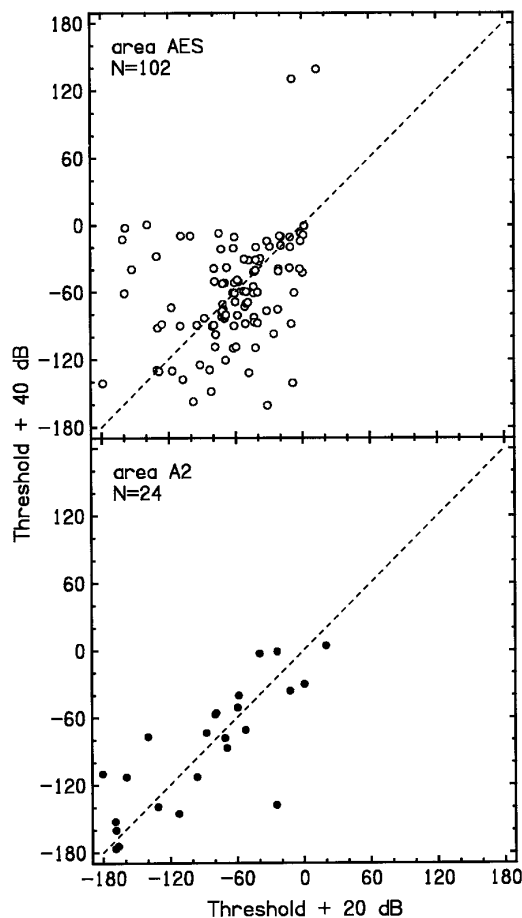
FIG. 6. Best-azimuth–centroid locations measured at 2 sound levels. This plot represents only the units that had measurable best-azimuth centroids at sound levels of 20 and 40 dB above units' thresholds. *Top* and *bottom*: areas AES and A2, respectively.

beginnings of penetrations, followed by a progression toward the contralateral pole, ending with a return to the frontal midline. That pattern, however, disappeared at sound levels 40 dB above threshold.

Artificial-neural–network recognition of spike patterns

Units responded to the onset of a noise burst with a burst of spikes, typically lasting 10–40 ms. For instance, the unit represented by the raster plot in Fig. 9 produced a burst of spikes lasting ~20–40 ms in response to noise bursts lasting 100 ms. Visual inspection of the raster plots shows that changes in stimulus azimuth resulted in changes in unit spike counts, latencies, and temporal dispersion of spikes. The conventional practice of representing unit responses simply by their spike counts potentially eliminates stimulus-related information that might be carried by the distribution of spikes in time. We wished to characterize neural responses in a way that would require a minimum of assumptions about how information is carried by a spike train. For that reason, we explored methods for recognizing complete spike patterns that contained both the magnitude of the neural response and the timing of spikes. As a measure of the stimulus-related information contained in spike patterns, we measured the accuracy with which we could identify stimulus

locations by recognizing the spike patterns elicited from particular locations.

The most suitable recognition algorithm that we found was an artificial neural network, trained with back-propagation (Rumelhart et al. 1986). The neural network was chosen because it is an effective general-purpose pattern recognizer, not because it models any particular biological structure. We used the training set of responses from each unit to train the network, with feedback from the associated stimulus azimuths, then presented the test set to the trained network and recorded the network's estimates of stimulus azimuths. The network architecture is schematized in Fig. 10. It consisted of a layer of four hidden units, which had hyperbolic tangent (i.e., nonlinear) transfer functions, and a layer of two linear output units. In initial work (e.g., Middlebrooks et al. 1994) we used a linear network, but we later found that a nonlinear network provided more accurate pattern recognition under conditions of varying stimulus intensity and under other conditions that we are examining in ongoing work. The input to the network consisted of bootstrap estimates of spike density functions, quantized in 1-ms bins. The 1-ms temporal resolution was chosen empirically. Resolution much coarser than 1 ms resulted in degradation in recognition performance, and finer resolution increased computation time without appreciable improvement in recognition performance. Similarly, the number of hidden units,
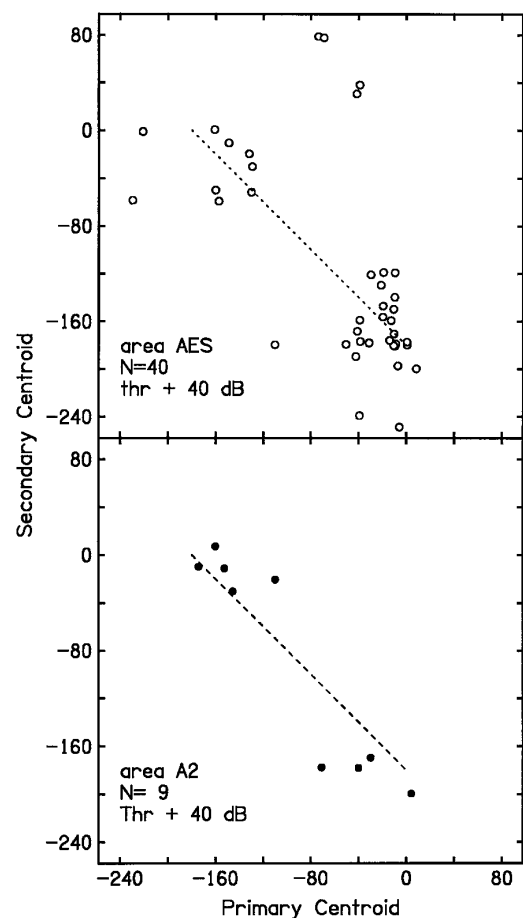

FIG. 7. Relation of primary and secondary centroids. These plots represent only units that had ≥2 peaks in their azimuth profiles. – – –, loci of points representing pairs of centroids that were located symmetrically with respect to the interaural axis. *Top* and *bottom*: areas AES and A2.
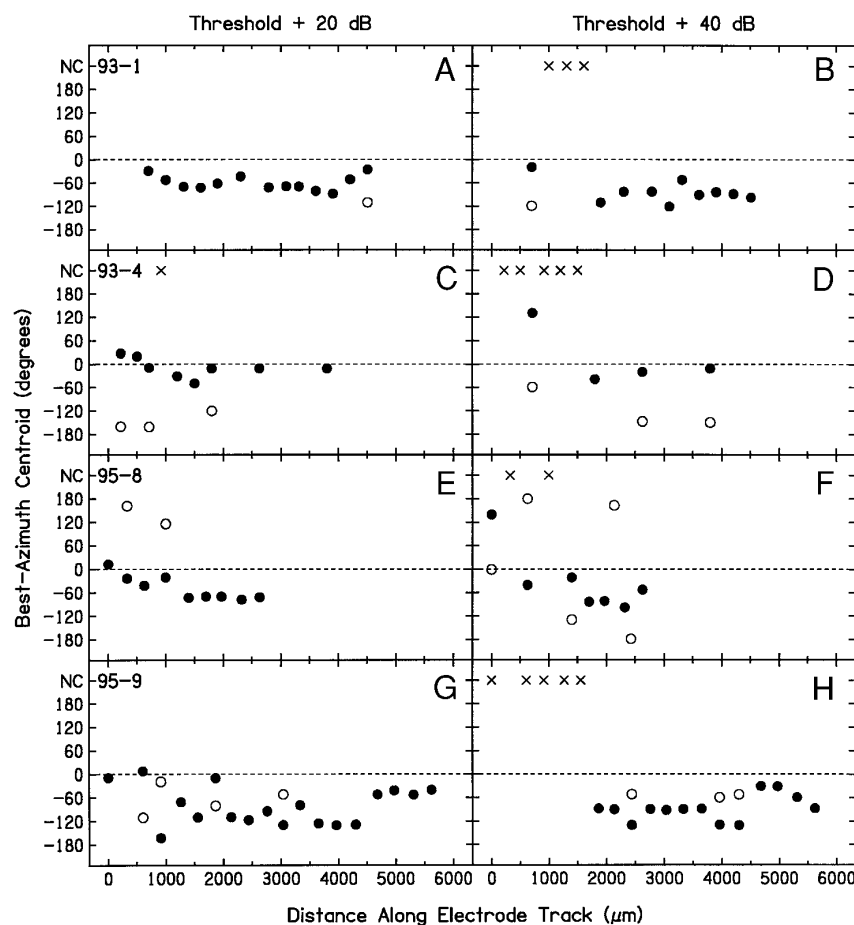
FIG. 8. Changes in centroid azimuth location with distance along electrode tracks. Four electrode tracks are represented, each with responses to sound levels 20 dB (*left*) and 40 dB (*right*) above units' thresholds. ●, primary centroids; ○, secondary centroids. Crosses represent units that showed <50% spike-count modulation and for which no centroid (NC) could be measured.

four, was chosen to optimize network performance across a large sample of units. The two output units were trained to estimate the sine and cosine of the sound-source azimuth, then the arctangent function was applied to produce an out-put in degrees of azimuth. This approach was chosen, rather than simply configuring the network to output azimuth directly, to avoid computational difficulties that resulted from the discontinuity in azimuth labels across the rear midline, where azimuths abruptly change from +179° to −180°.

The responses of units that we studied typically were rather sparse, so that many response patterns consisted of no more than one or two spikes. Average network classification of individual spike patterns often was little better than the level expected from random chance. For that reason, we chose to estimate spike density functions by averaging response patterns across trials. One approach would have been to form one average of all the responses in the training set for each stimulus location and one for each of the test-set responses. That approach, however, tended to cause over-training to the training set, and it provided only one instance of the test set for the purpose of evaluating performance. We adopted an alternative procedure, in which we formed multiple bootstrapped estimates of spike density functions (see METHODS).
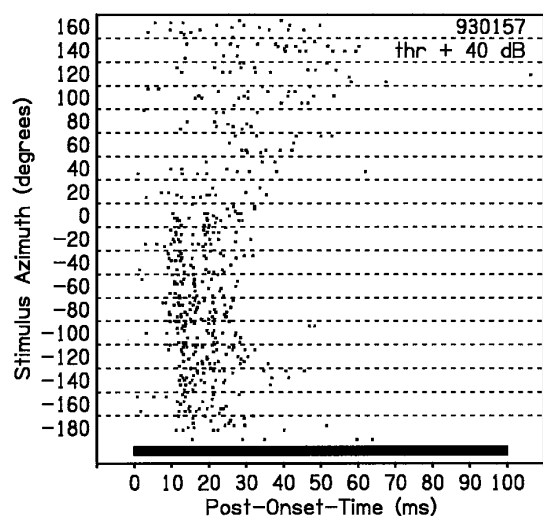
ESTIMATION OF STIMULUS LOCATIONS. The performance of an artificial neural network in estimating stimulus azimuth from the responses of one neuron is shown in Fig. 11; these results are from the unit represented by the raster plot in Fig. 9. In Fig. 11*A*, each plus represents the network estimate of azimuth based on one bootstrapped spike pattern, and solid line indicates the mean direction of responses at each stimulus azimuth. The dashed line with positive slope indi-
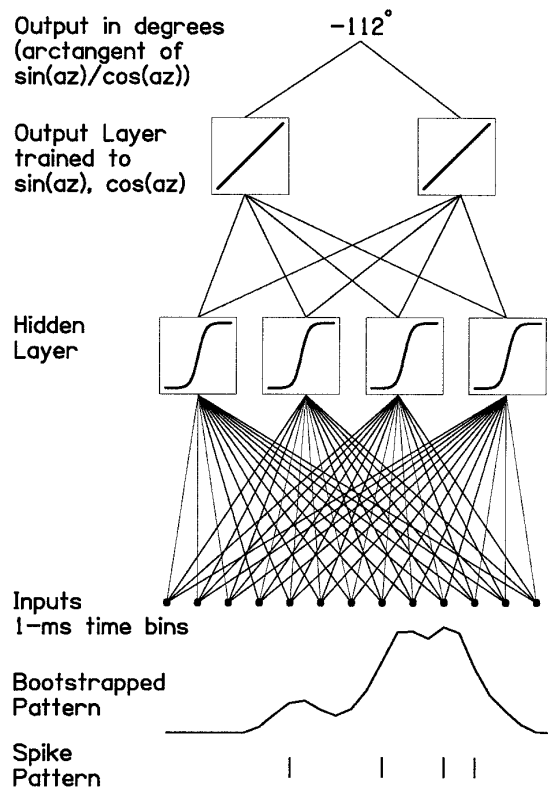

FIG. 9. Responses of *unit 930157*. This raster plot shows the responses of a unit to 100-ms noise bursts presented at various azimuths. ●, 1 spike from the unit. Each row of dots represents the spike pattern in response to 1 stimulus presentation. Stimulus azimuths were varied randomly, but in this plot responses are sorted according to stimulus azimuth as indicated on the vertical axis. Eight trials at each azimuth are represented. ■, duration of the stimulus.

FIG. 10. Artificial-neural–network architecture. Input to the network consisted of spike density functions that were averaged across 8 stimulus presentations and expressed in 1-ms time bins. Four units in the hidden layer had hyperbolic tangent transfer functions. Two units in the output layer had linear transfer functions. Network was feed-forward and fully connected. Network was trained with supervision so that the output units estimated the sine and cosine of the stimulus azimuth. Two outputs were represented with a single term by forming the arctangent of the 2 outputs.

cates perfect performance, and the dotted line with negative slope represent perfect front/back confusions, as in Fig. 7. Across azimuth, one can see some variation in the accuracy with which the mean direction matches the perfect performance line and considerable variation in the scatter of points around the mean directions. Nevertheless, it is noteworthy that the responses of this unit appeared to carry information about sound-source azimuth throughout 360° of azimuth. The responses of this unit distinguished stimulus left from right almost perfectly, as indicated by the near-absence of data points in the *top left* or *bottom right* quadrants of the figure, and it rarely confused front and back.

The distribution of errors in estimated azimuth for this unit is shown in Fig. 11B. The errors are binned with 20° resolution, which corresponds to the separation of the original sound-source locations. The expected value of each bin, given chance performance and 18 possible loudspeakers, would be 5.6%. In contrast, for unit in the figure, 21.8% of network errors were <10° in magnitude, indicating that the network assigned 21.8% of the bootstrapped spike patterns to the correct loudspeakers. We use the median magnitude of error, across all stimulus locations, as an overall measure of accuracy of azimuth coding by each unit. The median error is influenced by the mean direction of network estimates as well as by the scatter of estimates about the mean. In theory, the expected median error under conditions of random chance would be 90°. We tested our network algo-

rithm in a control condition in which the correspondence between spike patterns and stimulus locations was randomized. Across all units, the average of median errors in that condition was 86.8°. Presumably, the slight difference between that value and the predicted value of 90° indicates the
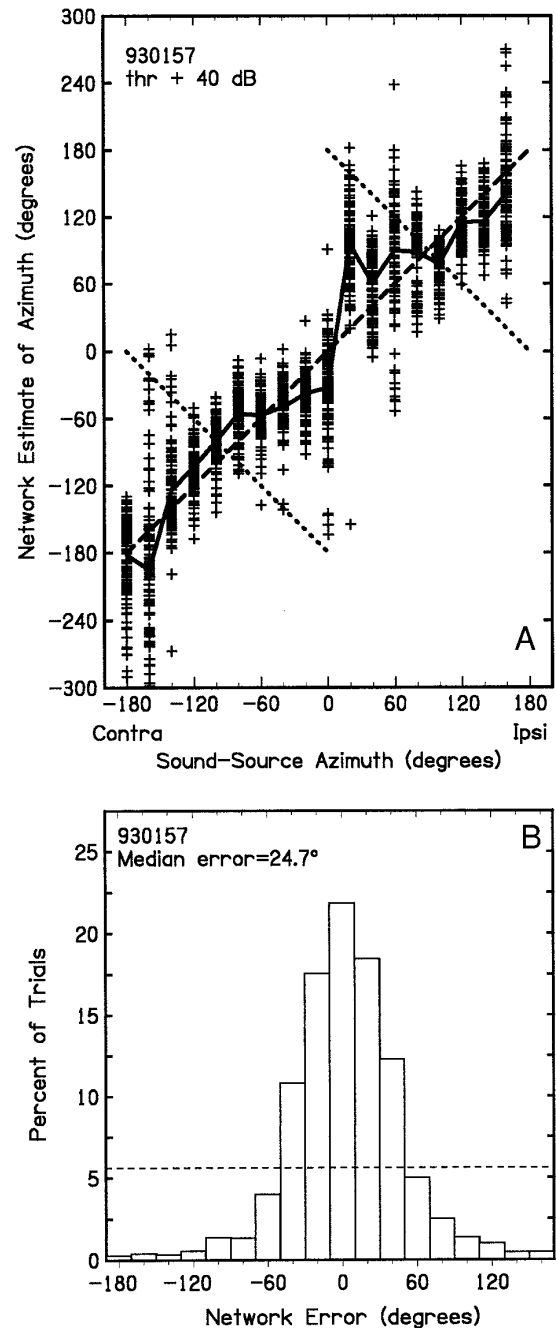

FIG. 11. Network performance for *unit 930157. A*: each + represents the network output in response to input of 1 bootstrapped pattern. Abscissa represents the actual stimulus azimuth and the ordinate represents the network estimate of azimuth. ——, mean directions of network estimates for each actual stimulus location. – – – (with positive slope), perfect performance; and · · · (with negative slope), front/back symmetry between stimulus location and network output. *B*: distribution of network errors. Bar at 0° indicates that the network output was within ±10° of the correct stimulus location on 21.8% of trials. – – –, 5.6%, which is the expected percentage of trials in each bin, given random chance performance and 18 bins. This unit produced a median error of 24.7°, which was among the best performances in our sample.

network's ability to exploit random variation in the outcome of training. The median error for the unit shown in Fig. 11 was 24.7°, which was among the lowest median errors in our sample.

Other units that showed relatively small median errors are represented in Fig. 12 by the mean directions of their network output. Local variations in the slope of the mean direction lines indicate that some locations were discriminated more accurately than were others. For instance, the responses of *unit 950902* discriminated among most locations throughout the contralateral and ipsilateral hemifields, with some errors around the front and rear midlines. In contrast, the responses of *unit 930142* failed to discriminate among locations within the ipsilateral hemifield, as indicated by the flat slope of the mean direction plot across the positive sound-source azimuths. For most of the units in our sample, the plots of mean direction showed an increase in slope near the frontal midline, indicating that changes in unit spike patterns with azimuth tended to be greatest across the midline.

Responses of the unit that is represented in Figs. 13 and 14 produced a median localization error of 45.7°, which is
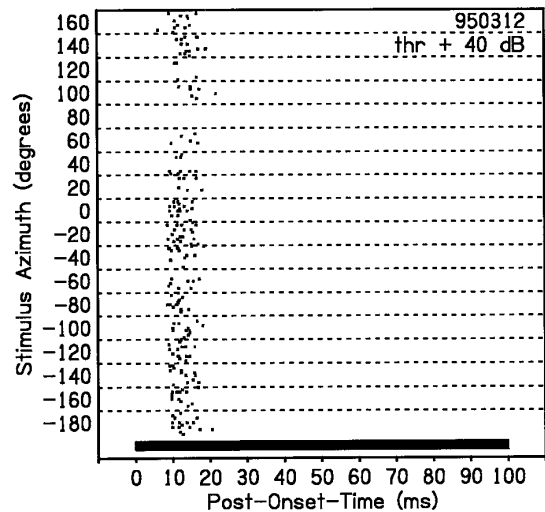
FIG. 13. Responses of *unit 950312*. Conventions are the same as in Fig. 9.

slightly larger than the mean of the sample for stimuli 40 dB above threshold. The responses of this unit discriminated reliably between contralateral and ipsilateral locations, as indicated by the abrupt jump in mean directions between stimuli at 0° and ipsilateral 20° and by the separation of the network responses at those two stimulus azimuths. With few exceptions, however, it failed to discriminate among locations within each sound hemifield. This unit was typical of many in that the network tended to assign most responses to one of two locations. In the distribution of errors shown in Fig. 14*B,* one can see that the network selected the correct speaker at greater than double the rate predicted by chance. The raster plot for this unit (Fig. 13) shows that its response was less sustained than that of the unit represented in Fig. 9, but across our sample of units, there was essentially no correlation between the duration of spike patterns and sizes of median errors ($r = 0.04$).

The accuracy of azimuth coding varied widely across our sample. We examined every unit that responded to noise bursts and that could be recorded long enough for complete study. That is, there was no selection of units on the basis of azimuth sensitivity. Figure 15 shows the distribution of median errors obtained for units in areas AES and A2 tested at sound levels 20 and 40 dB above units' thresholds. At each sound level, the means of the distributions for areas AES and A2 were not significantly different (*t*-test, $P > 0.05$). The modes of the distributions were near 35°. A median error of 35° indicates that half of the estimates of stimulus azimuth, based only on averages of eight responses of a single unit, fell within 35° of the actual stimulus azimuth. The median errors of all units were well below the chance level of 86.8°, and more than half of the median errors for either cortical area or for either sound level were smaller than half of the chance level.

Estimation of azimuth improved with increasing number of averages in the bootstrapped patterns. The median errors of network performance are shown for six units as a function of the number of averages in the training and test sets (Fig. 16); these six units are representative of the range of performance in our sample. Classifications of single spike patterns (i.e., "averages" of 1) tended to be rather inaccurate, but
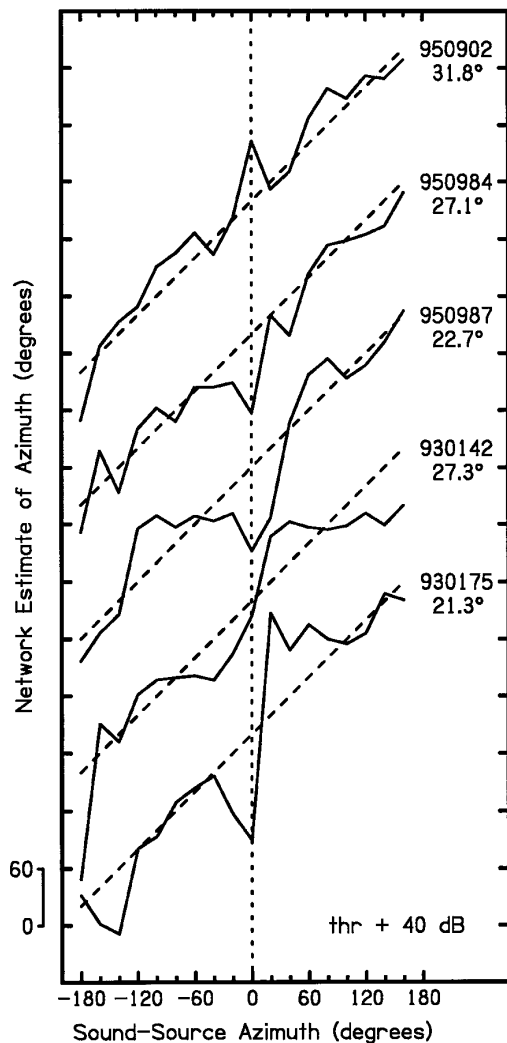

FIG. 12. Mean directions of network estimates for 5 units. Each solid line represents the mean direction of network estimates of azimuth for 1 unit. Unit number and median error is given next to each line. Sound levels were constant at 40 dB above each unit's threshold.
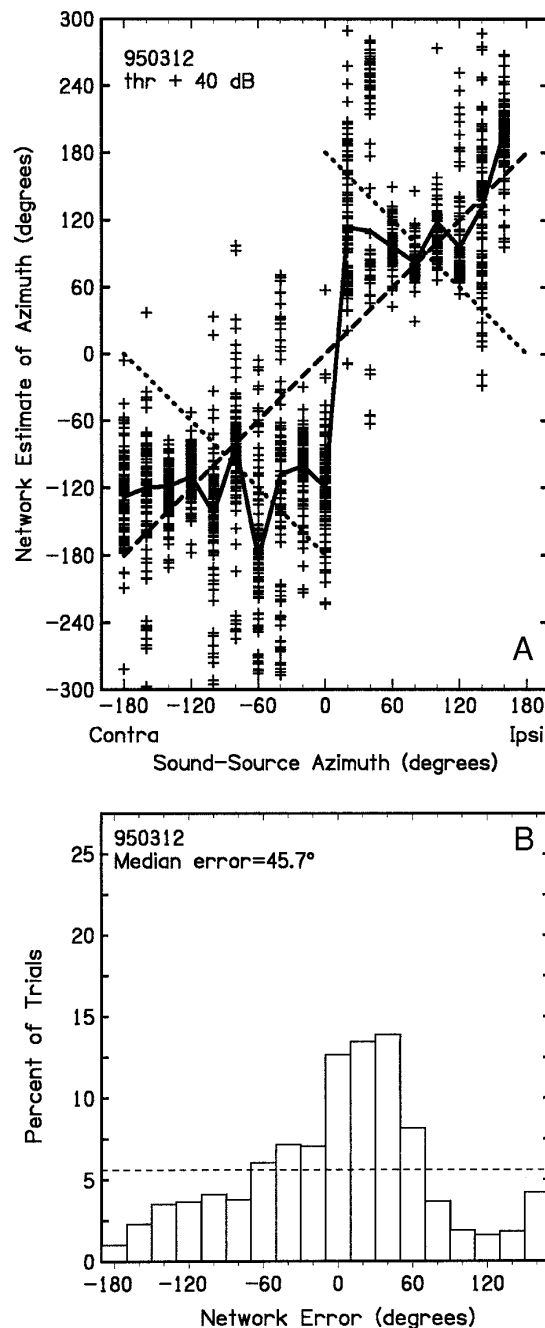
FIG. 14. Network performance for *unit 950312*. This unit produced a median error of 45.7°, which was slightly larger than the mean across our sample. Other conventions are the same as in Fig. 11.

median errors consistently were better than chance. Median errors for each of the units showed a steady decrease as numbers of averages increased to ~32 or 64, then tended to level off as the number of averages was increased further. With the exception of this figure, all the analysis presented in this report was based on averages of eight responses. Averages of eight provided sufficient spikes to permit recognition of spike timing yet were compatible with the numbers of training and test trials that we could record given the time limitations of data acquisition.

INFLUENCE OF SPL ON LOCATION CODING BY SPIKE PATTERNS. As considered in a previous section, the spatial tun-

ing of spike counts tended to change with changes in stimulus SPLs, so it is not surprising that spike patterns also changed with level changes. Both in area AES and in area A2, the average values of median errors in network output increased by ~6° as sound levels were increased from 20 to 40 dB above units' thresholds (Fig. 15). This is consistent with the observation that spike-count tuning for azimuth tended to broaden at higher SPLs. We tested a condition in which we trained a network with responses to one sound level, then tested with responses to another level. In that condition, median errors increased by averages of 7.3–17.5° (depending on cortical area and SPL) relative to the condition in which the network was trained and tested with responses to the same sound level. Despite the changes in spike patterns with changes in sound level, the networks could successfully recognize responses to test sets that varied in level between discrete steps of 20 and 40 dB above threshold when the networks were trained with training sets that similarly varied in level. We were concerned that the networks' success at recognizing test sets that varied in level might indicate that the network was in some way acquiring two level-specific maps of azimuth. That concern was alleviated by the results of two further tests. First, we predicted that limitations in the learning capacity of networks would limit the number of level-specific maps that a network could acquire, so increases in the number of sound levels should result in decreases in the accuracy of network performance. We trained and tested the network with responses to stimuli at five levels in 5-dB steps from 20 to 40 dB above threshold. Contrary to our prediction, performance under the five-level condition was significantly better than under the two-level condition. Second, we trained networks with 20- and 40-dB training sets, then tested the recognition of 30-dB test sets. Recognition of the 30-dB test set was nearly as accurate as when the networks were trained with 30-dB training sets. These results support the conclusion that, despite prominent level-related changes in spike patterns, artificial neural networks are capable of identifying azimuth-related features of spike patterns that are invariant across sound levels. The mean performance of networks under various conditions of sound level is summarized in Fig. 17.

INFLUENCE OF STIMULUS DURATION. One might predict that the response patterns of neurons would change with subtle changes in the envelopes of stimuli. We tested the sensitivity of response patterns of 57 units to rather extreme changes in stimulus envelopes by varying the overall durations of noise bursts from 1 to 100 ms. For each unit, sound levels were adjusted to constant levels relative to the threshold for each duration. Despite the large change in stimulus envelopes, the time course of response patterns and their dependence on stimulus azimuth were largely insensitive to stimulus durations. An example of the responses of one unit to stimuli of 1- and 100-ms durations is shown in Fig. 18. The accuracy of network estimation of azimuth varied somewhat among durations and among units, but the averages of median errors were not significantly different between durations of 1 and 100 ms ($n = 57$; $P > 0.05$, paired *t*-test, 2-tailed). When we tested spike patterns from 1-ms stimuli on networks that were trained with 100-ms training sets, median errors averaged only 6.7° larger than when same responses were tested on networks trained with 1-ms training sets. The
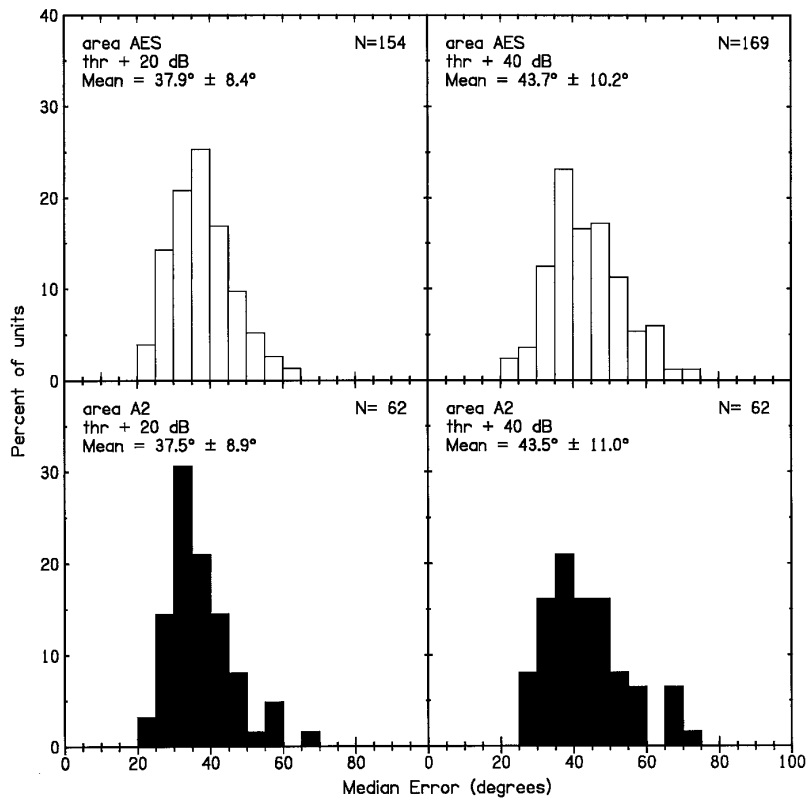
FIG. 15. Distributions of median errors. These histograms show the percentage of the units in each sample that produced network performances with particular ranges of median errors. *Top* and *bottom*: areas AES and A2, respectively. *Left* and *right*: 20 and 40 dB, respectively, above units' thresholds.

magnitude of those increases in median errors indicate that, although there was some degradation in recognition performance, recognition of responses was substantially retained across a 100-ms range of stimulus durations.

Panoramic coding of azimuth

A remarkable aspect of our analysis of azimuth coding by spike patterns is that single neurons appear to code locations
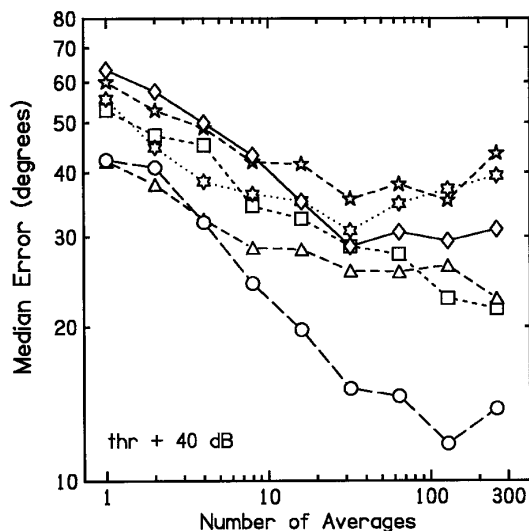

FIG. 16. Influence of number of averages on network performance. This plot shows the performance in network classification of averaged spike density functions from 6 units recorded in area AES. Each unit is represented by a different symbol. Bootstrap averages were formed that incorporated 1 to 256 samples, with replacement, from training sets that consisted of unit responses to 20 trials. Separate averages were formed from test sets that consisted of 20 additional responses.

throughout 360° of azimuth. This result is somewhat perplexing, given that the deficits in sound localization behavior that result from unilateral cortical lesions tend to be restricted to the side contralateral to the lesion. We measured azimuth coding contra- and ipsilateral to recording sites by training networks with responses to sounds distributed throughout 360°, then computing the median errors in recognition of responses to contra- versus ipsilateral stimuli; for the purpose of this analysis, we excluded locations on the midline (i.e., 0 and 180°). Localization performance was roughly equal on the two sides: the average of median errors was smaller for contralateral stimuli for 53% of units (81/154) in the 20/40-dB roving level condition and for 45% of units (70/154) in the 40-dB fixed level condition.

We were concerned that the balance of median errors between contra- and ipsilateral hemispheres might have resulted in some way from the design of the neural network algorithm that we used. For that reason, we used a second, independent measure to analyze responses to contralateral and ipsilateral stimuli. We used an ANOVA procedure to compare the amounts of variance in spike patterns that were accounted for by azimuth within the contralateral sound hemifield versus within the ipsilateral hemifield. That analysis confirmed that the spike patterns of many units could discriminate among ipsilateral azimuths as well or better than they could discriminate among contralateral azimuths. In the 20/40-dB roving-level condition, for example, stimulus azimuth accounted for a greater proportion of the total variance on the side ipsilateral to the recording site than on the contralateral side for slightly more than half of the units (59/108).

Azimuth coding by spike patterns and by spike counts

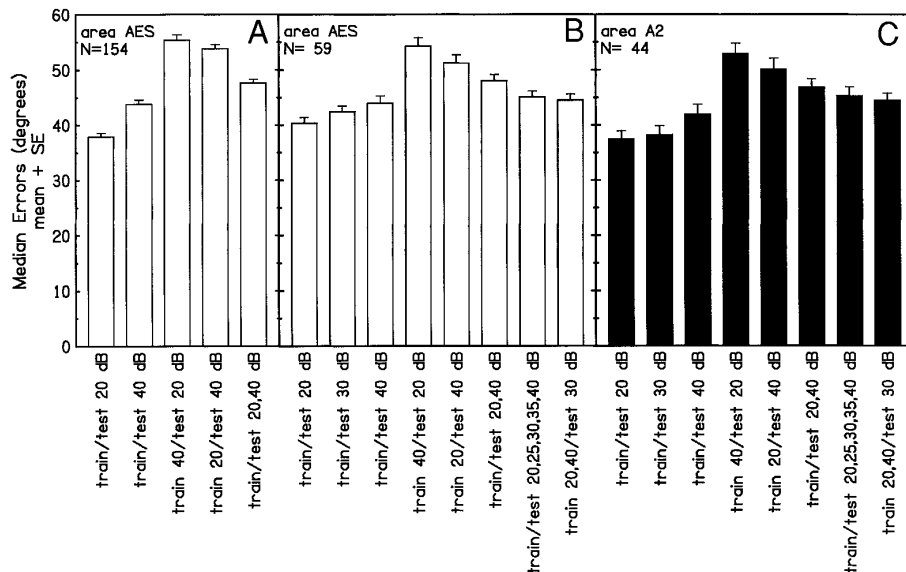We tested the hypothesis that coding of sound-source azimuth by spike patterns is more accurate than coding by spike

FIG. 17. Averages of median errors under various conditions of sound pressure. Each bar represents the mean ± SE of the mean of the median errors in azimuth localization under a condition in which the training and test sets consisted of responses to sounds at the stated levels above units' thresholds. For instance, "train 40/test 20 dB" indicates that the training sets consisted of responses to stimuli that were 40 dB above unit thresholds and that the test sets consisted of responses to stimuli that were 20 dB above thresholds. □, data from area AES; ■, data from area A2.

counts alone. The rationale for that hypothesis is that spike patterns contain all the information that is present in spike counts plus any additional information that might be available from the timing of spikes. We evaluated two methods

FIG. 18. Responses of *unit 960403* to noise bursts of 2 durations. *Top* and *bottom*: response to noises bursts that were 100 or 1 ms, respectively, in duration. Other conventions are as in Fig. 9.

to test azimuth coding by spike counts. The first method used artificial neural networks, as diagrammed in Fig. 10, except that the input consisted of one-dimensional spike counts instead of multidimensional spike patterns. The second method used a maximum-likelihood classifier. The advantage of the maximum-likelihood classifier was that it may be shown to be an optimal classifier in this one-dimensional situation (Green and Swets 1966; Neyman and Pearson 1933). In our situation, that means that it was optimal for identifying which of the 18 sound sources emitted the stimulus on each trial. In our sample of units, we compared the percentage of trials in which the maximum-likelihood classifier identified the correct loudspeaker with the percentage of trials in which the neural-network identification of spike counts produced an output within 10° of the correct loudspeaker location. As expected, the maximum likelihood classifier generally performed more accurately in that task. The disadvantage of the maximum likelihood classifier was that its output was quantized to particular loudspeaker locations and, thus was difficult to compare with the continuous output of the network analysis of spike patterns. Another disadvantage was that, when the maximum likelihood classifier produced an incorrect result, the error often would be quite large, whereas errors by the network procedure tended to scatter around the correct loudspeaker location. In all SPL conditions, the median errors produced by the maximum likelihood procedure were significantly larger than those produced by the network procedure ($P < 0.01$, all SPL conditions); for instance, median errors in the 20-dB fixed-level condition, averaged across 154 AES units, were 60.7° for the maximum likelihood classifier compared with 48.1° for the neural network. For that reason, we chose to use the network identification of spike counts to compare with network identification of spike patterns.

Figure 19 shows a comparison of azimuth coding by complete spike patterns with coding by spike counts, both classified by the neural network procedure. Data are from the 40-dB fixed-level condition. A considerable number of points in the two plots lie near the lines that indicated equal median errors. For the units represented by those points, the spike count apparently captured all the stimulus-related informa-
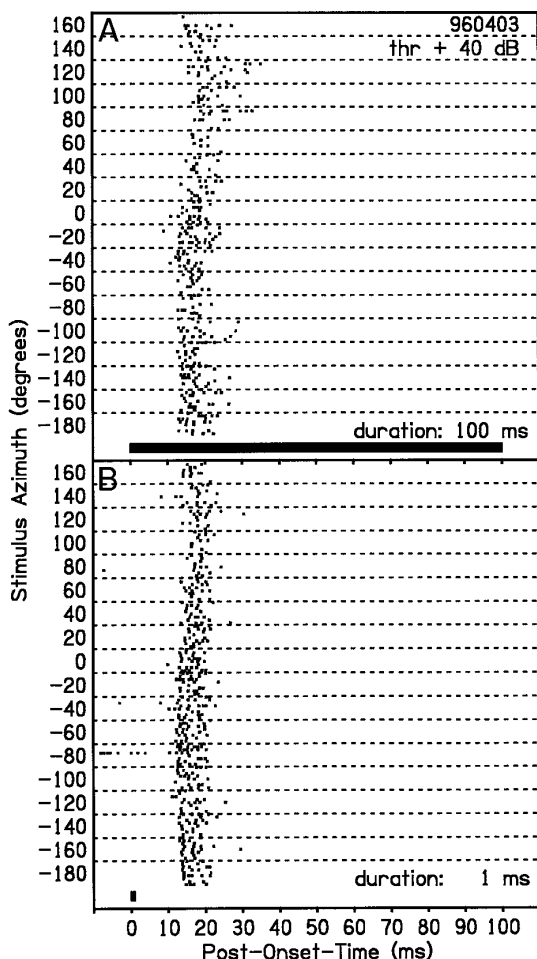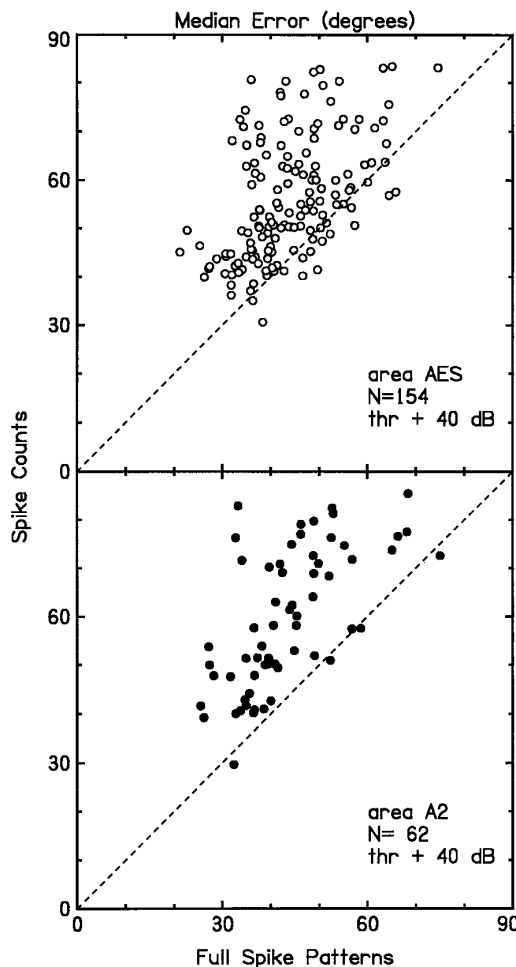
FIG. 19. Accuracy of azimuth coding by spike counts and by complete spike patterns. This plot shows the accuracy of artificial-neural–network estimation of sound-source azimuth based on full spike patterns and spike counts. Full patterns (abscissa) consisted of spike density functions expressed with 1-ms resolution. Spike counts (ordinate) were the total number of spikes in each density function; i.e., the area under the density function. *Top* and *bottom*: areas AES and A2.

DISCUSSION

In this study, we have explored two hypothetical codes by which neurons in the auditory cortex might represent the location of a sound source. We will refer to these here as a *topographic code* and a *distributed code*. We will consider in this DISCUSSION the implications of our data with regard to these codes. We also will consider the issue of information coding by the timing of neuronal spikes and will address some specific issues relating to area AES and the superior colliculus.

Topographical coding by tuned neurons

One can find many examples in the nervous system of neurons that are ''tuned'' in the sense that the neuron responds maximally or with lowest threshold to a particular stimulus feature or a particular value of a stimulus parameter. In some cases, the neuronal tuning can be attributed to the organization of the sensory periphery. For instance, the frequency tuning of neurons in the auditory cortex can be traced to the frequency analysis that is performed by the cochlea. In other cases, the neuronal tuning emerges from the integrative activity of the CNS (e.g., Knudsen et al. 1987). Examples include the tuning for sound-source location that has been demonstrated in the superior colliculus and optic tectum (Knudsen 1982; Middlebrooks and Knudsen 1984; Palmer and King 1982) and the tuning for parameters of echolocation signals in the bat's auditory cortex (Suga 1990). Barlow (1972) formalized the notion of stimulus coding by tuned neurons in his ''neuron doctrine.'' Central to that doctrine was the notion that sensory neurons are tuned to specific ''trigger features'' and that a strong discharge by a neuron would signal the presence of a trigger feature within its receptive field. Barlow postulated that the consequence of this neuronal specificity is that a given stimulus would be represented by a minimum number of active neurons. A specific example from Barlow's work is the ''bug detector'' of the frog retina, a class of ganglion cells that respond with great specificity to small black disks moving within neurons' receptive fields (Barlow 1953; also see Lettvin et al. 1959). The notion of tuned neurons put forward by Barlow and others, and the demonstration of such neurons in many systems, has had a pervasive influence on sensory physiology.

Most previous studies of spatial coding in the auditory cortex have been designed around the hypotheses that cortical neurons are more or less sharply tuned for sound-source location and that the location of a sound source is mapped by the cortical location of a small population of maximally active neurons. Examples of such organization are found in certain noncortical structures, specifically the optic tectum in the barn owl (Knudsen 1982) and the mammalian superior colliculus (Palmer and King 1982). Published results from the auditory cortex, however, generally fail to support the hypothesis that sound location is represented by a systematic map constituted of sharply tuned neurons. For instance, one study of cortical area A1 (Middlebrooks and Pettigrew 1981) showed that although some neurons in area A1 had restricted receptive fields for stimuli at low SPL, the spatial tuning of most units broadened considerably as sound levels were increased more than ∼10 dB above units' thresholds. More recent studies of area A1 have shown that short se-

tion that was contained in the spike patterns. Nevertheless, the large majority of the points lie well above the equal-performance lines. In each condition of sound level (i.e., 20- dB fixed, 40-dB fixed, 2 roving levels, 5 roving levels) in areas AES and A2, median errors obtained with complete spike patterns averaged from 7.3 to 16.2° smaller than those obtained with spike counts alone (paired *t*-test, $P < 0.01$, all conditions). In the 40-dB condition that is illustrated, the spike counts of 19% of AES units (29/154) and 34% of A2 units (21/62) produced median errors of ≥70°, whereas only one unit in our sample from each area showed such near-chance performance with complete spike patterns. Conversely, the spike counts of only 5% of AES units and 3% of A2 units could produce median errors <40°, whereas the spike patterns of 40% of AES units and 45% of A2 units could produce such levels of accuracy. These data strongly support the hypothesis that, overall, azimuth coding by complete spike patterns is more accurate than azimuth coding by spike counts. The points that lie well above the equal-performance line represent units that presumably carry additional stimulus-related information in the timing of spikes.

quences of units recorded along electrode tracks through the cortex can exhibit spatial tuning that shifts systematically in azimuth according to shifts in cortical place (Clarey et al. 1994; Imig et al. 1990; Rajan et al. 1990b). Nevertheless, such sequences typically show reversals in the direction of azimuth-tuning shifts and often are interrupted by units that show broad azimuth tuning. Again the tuning of most units broadens considerably as SPLs are increased.

In our analysis of data from areas AES and A2, we evaluated several ways of representing the spatial preferences of neurons, and we settled on ''best-azimuth centroids'' as a measure that seemed most likely to reveal an auditory space map, if it was present. When stimuli were only 20 dB above unit thresholds, we sometimes observed segments of electrode tracks as long as 1.5 mm in which units showed systematic shifts in the locations of their best-azimuth centroids. The lengths of these electrode-track segments were comparable with the longest such sequences of monotonically shifting tuning that have been reported from studies of area A1. When stimulus SPLs were increased by another 20 dB, however, the centroids of many units tended to shift in location, and the modulation depth of other units decreased to the point that centroids could no longer be measured. Thus in our results from area AES and A2 and in published results from A1, there is some evidence for short, discontinuous maps of azimuth based on single-unit tuning. Those maps degrade, however, as sound levels are raised to moderate levels, i.e., sound-source levels that most human listeners would regard as comfortable and easily localizable.

Our measurements of the azimuth tuning of units permit us to infer the size of the population of activated neurons in the sampled regions of areas AES and A2 as a function of sound-source level and azimuth. It appears that a sound-source presented at a moderate SPL at nearly any azimuth in the horizontal plane would activate nearly all auditory units in areas AES and A2 of both hemispheres to >25% of their maximum rates. Most of the auditory units on the side contralateral to the sound source would be activated to greater than half of their maximum rates. This inference is very different from the usual topographic model, which predicts that a particular trigger feature of a stimulus is represented by a restricted population of active neurons. Although the spatial tuning that one can record in areas AES and A2 must result from the processes that subserve sound localization, it is difficult to accept that topographic maps constituted of spatially tuned units are the principal representation of sound-source location.

One possible explanation for why we have failed to find a cortical map of auditory space is that we might have examined the wrong cortical area. We chose to study nontonotopic areas AES and A2 because neurons there have broad frequency bandwidths and because accurate sound localization requires integration of information across frequencies (see Middlebrooks and Green 1992). We were interested particularly in area AES because it sends descending projections to the space-mapped superior colliculus (Meredith and Clemo 1989), although certain characteristics of that anatomic projection, considered in a later section, are not compatible with the existence of a topographic map in area AES. One might argue that area A1 would have been a better area in which to search for a spatial representation because experimental lesions restricted to area A1 have been shown to produce

behavioral deficits in localization (Jenkins and Merzenich 1984). Published studies of A1, however, have concluded that there is no continuous map of space in A1 (Clarey et al. 1994; Middlebrooks and Pettigrew 1981; Rajan et al. 1990a). Similarly, a study of spatial sensitivity in the anterior auditory area revealed neither a map nor spatial tuning that was noticeably sharper than has been seen in areas AES, A2, or A1 (Korte and Rauschecker 1993). We studied only the central part of area A2, so we would have missed an auditory space map that might have been hidden on the bank of the posterior ectosylvian sulcus, or there might be an as-yet-undiscovered space map in the posterior or ventral posterior auditory fields. An alternative explanation, however, is that the hypothesis of a topographic code is simply not correct in the case of auditory spatial coding. For that reason, we have begun to explore alternative forms of cortical representation of location.

Distributed coding by panoramic neurons

We have tested a distributed model in which the activity of individual units is assumed to code, with varying levels of accuracy, the locations of sound sources throughout 360° of azimuth. That is, individual units are assumed to be panoramic. This model is inspired by the work of Bialek and colleagues (1991), who showed that one could ''read'' the direction of whole-field visual motion from the firing pattern of a single neuron in the visual system of a fly. In a somewhat analogous way, we could read the locations of sound sources from the firing patterns of single cortical neurons. We obtained a moderate level of localization performance based on information carried by spike counts, and that performance improves when we incorporate the timing of spikes within spike patterns. Even the units in our sample that showed the worst performance could distinguish left from right with a reasonable level of accuracy. The units that showed the best performance could accurately distinguish left from right, front from back, and could discriminate among locations within each of the four quadrants.

The distribution of median errors in our sample was unimodal, ranging continuously from the smallest to the largest. For that reason, we saw no justification for identifying particular subpopulations of location-specific or nonspecific units. Nevertheless, one might imagine that sound localization is accomplished by information that is distributed among the units that showed the smallest median errors and that the units that showed large median errors have little or no role in location coding.

Localization by the best units, with median errors <25°, still was much worse than that localization performance of a behaving animal. It is important to realize, however, that our test of location-coding accuracy incorporated the responses of only one cortical neuron at a time, whereas the pool of neurons that contribute to a behaving cat's localization judgment presumably is several orders of magnitude larger. Our practice of averaging spike patterns across multiple trials is a very simplistic way of simulating the information that is carried by multiple neurons, and we found that localization performance improved with the number of averages (i.e., Fig. 16). In a more realistic situation in which information actually is contributed by multiple neurons, we presume that information would be added both by increases

in the number of independent samples and by information that would be carried by the relative activity among neurons.

A critical factor that argues against a topographic model of spatial coding in the auditory cortex is that the spatial tuning of most neurons tends to broaden considerably with increases in stimulus SPL. As spatial tuning broadens, the spike patterns elicited at a particular location also change. For that reason, an artificial neural network that is trained with responses to one sound level typically performs poorly in recognizing the responses to a different level. Nevertheless, networks could be trained to recognize the responses of stimuli that varied in level, and the performance in the roving level condition improved as the number of increments in level was increased. The ability of a network to estimate the locations of sources that vary in SPL suggests that one or more features of spike patterns are somewhat invariant across a range of levels, even though other features such as spike count tend to vary with SPL. This speculation is consistent with work in the primate visual cortex. There, Gawne and colleagues (1996) find that spike counts tend to code stimulus orientation, whereas spike latencies code stimulus contrast, and Victor and Purpura (1996) find that visual contrast and visual texture are coded by properties of spike patterns that differ in time scale by nearly an order of magnitude.

Generally, an artificial neural network that was trained to recognize the responses of one unit could not accurately classify the responses of another unit. This is not surprising because even such a basic response feature as the spike count varies between units. Even as an artificial neural network must be custom trained for each neuron, however, one might expect that the network of neurons to which a given cortical neuron projects also is custom trained to the responses of that neuron. Although the basic rules for recognizing spike patterns, such as the number and timing of spikes, presumably are constant between neurons, there is no reason to assume that the actual details of spike patterns are constant among neurons.

We found that the responses of single neurons appear to carry location-related information for locations throughout 360° of azimuth. This result is somewhat contrary to expectations based on results from ablation-behavioral experiments in which a unilateral cortical lesion produced a sound-localization deficit restricted to the side contralateral to the lesion (Jenkins and Masterton 1982). We tested the hypothesis that, although location coding by single neurons appears to be panoramic, coding accuracy might be better for contralateral locations. That hypothesis proved to be incorrect because tests of neural-network performance show about equal accuracy in both hemispheres. An alternative hypothesis is that the behavior of the animal is dominated by units that are strongly activated. Generally, we found that neurons tended to respond with more spikes to sounds in the contralateral hemifield compared with the ipsilateral hemifield. Nevertheless, our network design permitted recognition of relatively weak responses. For instance, even when a stimulus resulted in no spikes, the network would produce a specific output, generally in the ipsilateral hemifield. Perhaps an animal that sustains a unilateral cortical lesion is not able to derive location information from the weak responses that are evoked by a sound source that is ipsilateral to the surviving hemisphere.

Information carried by spike timing

For most units, sound-source locations could be estimated more accurately from complete spike patterns than from spike counts alone. A spike pattern contains a representation of every spike, so it contains any stimulus-related information that is carried by spike counts. The full pattern also contains the distribution of spikes in time, so presumably the superior localization performance obtained with complete spike patterns results from information that is carried by spike timing. The premise of our experiments is that if we can estimate sound-source locations by recognizing spike patterns, then the spike patterns must carry stimulus-related information. This is an empiric measure of information. In contrast, previous studies of the stimulus coding in the primate visual cortex have used an information theoretic approach. Richmond and Optican (1987, 1990) represented spike patterns as weighted sums of a small number of principal components, then assayed the amount of stimulus-related information that was carried by each component. The weight on the highest-valued component tended to correlate with unit spike counts, confirming one's intuition that the spike count is an important information-bearing feature of a spike pattern. The second- and lower-valued principal components also were shown to carry stimulus-related information. That observation suggests that features of spike patterns additional to spike counts, presumably spike timing, carry stimulus-related information in the visual cortex.

Brugge and colleagues (1996) have shown that the first-spike latencies of units in area A1 can vary over a range of ∼5 ms as the virtual location of a sound source is varied within the units' virtual spatial receptive fields. About half of their sample showed latencies that were ''ordered'' as a function of virtual location, whereas the other half showed latencies that were ''disordered.'' The A1 units that showed ordered latencies might be a counterpart of the units in areas AES and A2 that showed the largest improvement in network performance when tested with full spike patterns compared with spike counts alone.

Victor and Purpura (1996) used a metrical analysis of spike patterns in the visual cortex. They found that temporal features on a scale of 10–30 ms were significant for coding visual contrast and that features on a scale of 100 ms were important for coding visual texture. In our study, we found that accuracy of location coding deteriorated when the spike patterns were expressed with resolution coarser than ∼1 ms. The differences in time scales of coding in the auditory and visual cortices are consistent with the general organization of the auditory system to preserve temporal fidelity. For instance, synapses in the cochlear nucleus and superior olivary complex preserve stimulus phase information with high precision, such that human subjects can detect interaural delays on the order of tens of microseconds (Zwislocki and Feldman 1956). The entire response of a cat auditory cortical neuron to a noise burst typically consists of a latent period of ∼12 ms followed by a burst of spikes lasting 10–40 ms. In contrast, minimum latencies in the primate visual cortex are ∼35 ms (e.g., Gawne et al. 1996) and burst durations, even of phasic neurons, appear to last ≥50 ms (e.g., Richmond and Optican 1987).

Evidence that temporal features of spike patterns carry stimulus-related information does not constitute evidence

that such information actually influences perception and behavior. Indeed, the possible contribution of spike timing to stimulus coding in the visual cortex is a matter of some debate (e.g., Shadlen and Newsome 1994; Softky 1995). It is a challenging problem to show directly that any particular aspect of a neuron's response (even spike count or rate) influences behavior, although one can demonstrate substantial correlations between neural firing and behavior (e.g., Britten et al. 1992; Mountcastle et al. 1990). Indirectly, if a coding task could be performed by spike patterns but not by spike counts, that would support the hypothesis that temporal aspects of spike patterns are significant for behavior. Many of the units that we studied showed that property quantitatively, in that they showed reasonable accuracy in spatial coding with their spike patterns but near-chance performance with their spike counts. In future experiments, we hope to test coding tasks that might reveal qualitative differences in the coding capacity of particular features of spike patterns.

Relation to spatial coding in the superior colliculus

We chose area AES for this study of spatial coding because it is the only auditory cortical area that has been shown to project strongly to the superior colliculus (Meredith and Clemo 1989). In the superior colliculus of anesthetized cats, auditory neurons show spatial tuning, and the tuning of neurons varies as a function of location within the colliculus to form a map of auditory space (Middlebrooks and Knudsen 1984). Results of experiments in behaving primates suggest that it might be more appropriate to regard the auditory map in the superior colliculus as one of several components of a map of motor error (e.g., Sparks 1988). For that reason, one could argue that the particular mapped format of the auditory representation in the superior colliculus is imposed by the constraints of the overlying retinal and eye movement maps. In the projection from area AES, cells within a restricted region of area AES diverge to sites distributed throughout the map in the colliculus and, conversely, any given region of the auditory map in the colliculus receives projections that converge from sources distributed throughout area AES. This diffuse projection from area AES to the superior colliculus is not compatible with a topographic auditory map in area AES projecting point-to-point to a topographic map in the superior colliculus. It is compatible, however, with a model in which neurons distributed throughout area AES each contain information about locations throughout the sound field.

Jay and Sparks (1984) have shown that the auditory spatial tuning of neurons in the superior colliculus of awake monkeys is modulated by the position of the eyes in the orbits. A similar result has been obtained in the cat (Hartline 1995; Peck et al. 1995). It is difficult to imagine how this dynamic auditory spatial tuning could result from rigidly space-tuned inputs. It is an appealing notion that dynamic processes within the superior colliculus might modulate the way in which those neurons decode panoramic inputs from area AES.

Concluding remarks

In support of a topographic hypothesis of cortical spatial coding, we find that the spike counts of some neurons in areas AES and A2 show spatial tuning, at least at low sound levels. Moreover, the centroids of azimuth tuning, in some cases, can be seen to shift systematically as a function of unit location across as much as 1.5 mm of cortex. Also, reports that unilateral cortical lesions produce contralateral localization deficits indicate the presence of, at least, a macroscopic map from sound hemifield onto cortical hemisphere. Contrary to the topographic hypothesis, the spatial tuning of most units tends to broaden, and azimuth centroids can shift, as sound levels are increased to moderate levels. The topography that can be seen is fragmentary and tends to deteriorate at moderate sound levels. Although one cannot ignore the spatial tuning that one sees at low sound levels, it is difficult to conclude that the primary code for sound-source location involves single units that are specialized to signal the presence of a sound source within particular spatial receptive fields. In support of a distributed hypothesis, the spike patterns of individual units appear to carry at least partial information about the location of a sound-source anywhere in the sound field. That information appears to be coded by the number of spikes in a response pattern and by the timing of those spikes. Units with similar panoramic spatial properties are distributed throughout areas AES and A2. It remains to be seen whether the task of spatial coding is the function of specialized subpopulations of units or whether that function is shared among large populations of units that vary widely in their individual localization accuracy.

With the use of biologically plausible delay lines and coincidence detectors, one might devise a model of a biological neural network that could decode the spike patterns of cortical neurons, much as we attempted to decode spike patterns with artificial neural networks. Such a model reduced to a single neuron, however, would return us conceptually to the spatially selective neuron that has proven to be so elusive in the auditory cortex. As an alternative, we are exploring the possibility that there is no overt decoding of cortical spike patterns by single neurons and that, instead, the decoded output is apparent only in the behavior of the intact animal. One possibility is that the magnitude and timing of spike patterns of neurons might act to synchronize specific subpopulations of neurons, thereby selecting those populations to contribute to behavioral output. In ongoing experiments, we are exploring the principles by which stimuli might be coded by the coordinated activity within distributed populations of cortical neurons.

REFERENCES

AHISSAR, M., AHISSAR, E., BERGMAN, H., AND VAADIA, E. Encoding of sound-source location and movement: activity of single neurons and inter-

actions between adjacent neurons in the monkey auditory cortex. *J. Neurophysiol*. 67: 203–215, 1992.

BARLOW, H. B. Summation and inhibition in the frog's retina. *J. Physiol. (Lond.)*. 119: 69–88, 1953.

BARLOW, H. B. Single units and sensation: a neuron doctrine for perceptual psychology? *Perception* 1: 371–394, 1972.

BIALEK, W., RIEKE, F., STEVENINCK, R.R.D.R.V., AND WARLAND, D. Reading a neural code. *Science* 252: 1854–1857, 1991.

BRITTEN, K. H., SHADLEN, M. N., NEWSOME, W. T., AND MOVSHON, J. A. The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J. Neurosci.* 12: 4745–4765, 1992.

BRUGGE, J. F., REALE, R. A., HIND, J. E., CHAN, J.C.K., MUSICANT, A. D., AND POON, P.W.F. Simulation of free-field sound sources and its application to studies of cortical mechanisms of sound localization in the cat. *Hear. Res.* 73: 67–84, 1994.

BRUGGE, J. F., REALE, R. A., AND HIND, J. E. The structure of spatial receptive fields of neurons in primary auditory cortex of the cat. *J. Neurosci.* 16: 4420–4437, 1996.

CLAREY, J. C., BARONE, P., AND IMIG, T. J. Functional organization of sound direction and sound pressure level in primary auditory cortex of the cat. *J. Neurophysiol.* 72: 2383–2405, 1994.

DEMUTH, H. AND BEALE, M. *Neural Network Toolbox for Use with Matlab*. Natick, MA: The Mathworks, 1995.

EFRON, B. AND TIBSHIRANI, R. Statistical data analysis in the computer age. *Science* 253: 390–395, 1991.

FISHER, N. I., LEWIS, T. L., AND EMBLETON, B.J.J. *Statistical Analysis of Spherical Data*. New York: Cambridge, 1987.

GAWNE, T. J., KJAER, T. W., AND RICHMOND, B. J. Latency: another potential code for feature binding in striate cortex. *J. Neurophysiol.* 76: 1356–1360, 1996.

GREEN, D. M. AND SWETS, J. A. *Signal Detection Theory and Psychophysics*. New York: Wiley, 1966.

GREENE, T. C. The ability to localize sound: a study of binaural hearing in patients with tumor of the brain. *Arch. Surg.* 18: 1825–1841, 1929.

HARTLINE, P. H. Effects of eye position on auditory localization and neural representation of space in superior colliculus of cats. *Exp. Brain Res.* 104: 402–408, 1995.

HAYS, W. L. *Statistics*, New York: Holt Reinhart and Winston, 1981.

IMIG, T. J., IRONS, W. A., AND SAMSON, F. R. Single-unit selectivity to azimuthal direction and sound pressure level of noise bursts in cat high-frequency primary auditory cortex. *J. Neurophysiol.* 63: 1448–1466, 1990.

JAY, M. F. AND SPARKS, D. L. Auditory receptive fields in primate superior colliculus shift with changes in eye position. *Nature* 309: 345–347, 1984.

JENKINS, W. M. AND MASTERTON, R. B. Sound localization: effects of unilateral lesions in central auditory system. *J. Neurophysiol.* 47: 987–1016, 1982.

JENKINS, W. M. AND MERZENICH, M. M. Role of cat primary auditory cortex for sound-localization behavior. *J. Neurophysiol.* 52: 819–847, 1984.

KING, A. J. AND HUTCHINGS, M. E. Spatial responses properties of acoustically responsive neurons in the superior colliculus of the ferret: a map of auditory space. *J. Neurophysiol*. 57: 596–624, 1987.

KJAER, T. W., HERTZ, J. A., AND RICHMOND, B. J. Decoding cortical neuronal signals: network models, information estimation and spatial tuning. *J. Computat. Neurosci.* 1: 109–139, 1994.

KLINGON, G. H. AND BONTECOU, D. C. Localization in auditory space. *Neurology* 16: 879–886, 1966.

KNUDSEN, E. I. Auditory and visual maps of space in the optic tectum of the owl. *J. Neurosci.* 2: 1177–1194, 1982.

KNUDSEN, E. I., DU LAC, S., AND ESTERLY, S. D. Computational maps in the brain. *Annu. Rev. Neurosci.* 10: 41–65, 1987.

KORTE, M. AND RAUSCHECKER, J. P. Auditory spatial tuning of cortical neurons is sharpened in cats with early blindness. *J. Neurophysiol.* 70: 1717–1721, 1993.

LETTVIN, J. Y., MATURANA, H. R., McCULLOCH, W. S., AND PITTS, W. H. What the frog's eye tells the frog's brain. *Proc. IRE,* 1959.

MEREDITH, M. A. AND CLEMO, H. R. Auditory cortical projection from the anterior ectosylvian sulcus (field AES) to the superior colliculus in the cat: an anatomical and electrophysiological study. *J. Comp. Neurol.* 289: 687–707, 1989.

MIDDLEBROOKS, J. C., CLOCK, A. E., XU, L., AND GREEN, D. M. A pan-

oramic code for sound location by cortical neurons. *Science* 264: 842–844, 1994.

MIDDLEBROOKS, J. C. AND GREEN, D. M. Sound localization by human listeners. *Annu. Rev. Psychol.* 42: 135–159, 1991.

MIDDLEBROOKS, J. C. AND KNUDSEN, E. I. A neural code for auditory space in the cat's superior colliculus. *J. Neurosci.* 4: 2621–2634, 1984.

MIDDLEBROOKS, J. C. AND KNUDSEN, E. I. Changes in external ear position modify the spatial tuning of auditory units in the cat's superior colliculus. *J. Neurophysiol.* 57: 672–687, 1987.

MIDDLEBROOKS, J. C. AND PETTIGREW, J. D. Functional classes of neurons in primary auditory cortex of the cat distinguished by sensitivity to sound location. *J. Neurosci.* 1: 107–120, 1981.

MOUNTCASTLE, V. B., STEINMETZ, M. A., AND ROMO, R. Frequency discrimination in the sense of flutter: psychophysical measurements correlated with postcentral events in behaving monkeys. *J. Neurosci.* 10: 3032–3044, 1990.

NEYMAN, J. AND PEARSON, E. S. On the problem of the most efficient test of statistical hypotheses (Abstract). *Philos. Trans. R. Soc. Lond. A Math. Phys. Sci.* 231: 289, 1933.

PALMER, A. R. AND KING, A. J. The representation of auditory space in the mammalian superior colliculus. *Nature* 299: 248–249, 1982.

PECK, C. K., BARO, J. A., AND WARDER, S. M. Effects of eye position on saccadic eye movements and on the neuronal responses to auditory and visual stimuli in cat superior colliculus. *Exp. Brain Res.* 103: 227–242, 1995.

RAJAN, R., AITKIN, L. M., AND IRVINE, D.R.F. Azimuthal sensitivity of neurons in primary auditory cortex of cats. II. Organization along frequency-band strips. *J. Neurophysiol.* 64: 888–902, 1990a.

RAJAN, R., AITKIN, L. M., IRVINE, D.R.F., AND McKAY, J. Azimuthal sensitivity of neurons in primary auditory cortex of cats. I. Types of sensitivity and the effects of variations in stimulus parameters. *J. Neurophysiol.* 64: 872–887, 1990b.

RICHMOND, B. J. AND OPTICAN, L. M. Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. II. Quantification of response waveform. *J. Neurophysiol.* 57: 147–161, 1987.

RICHMOND, B. J. AND OPTICAN, L. Temporal encoding of two-dimensional patterns by single units in primary visual cortex. II. Information transmission. *J. Neurophysiol.* 64: 370–380, 1990.

RUMELHART, D. E., HINTON, G. E., AND WILLIAMS, R. J. Learning internal representations by error propogation. In: *Parallel Data Processing,* edited by D. E. Rumelhart and J. McClelland. Cambridge, MA: MIT Press, 1986, vol. 1, chap. 8, p. 318–362.

SANCHEZ-LONGO, L. P. AND FORSTER, F. M. Clinical significance of impairment of sound localization. *Neurology* 8: 119–125, 1958.

SHADLEN, M. N. AND NEWSOME, W. T. Noise, neural codes and cortical organization. *Curr. Opin. Neurobiol.* 4: 569–579, 1994.

SOFTKY, W. R. Simple codes versus efficient codes. *Curr. Opin. Neurobiol.* 5: 239–247, 1995.

SPARKS, D. L. Neural cartography: sensory and motor maps in the superior colliculus. *Brain Behav. Evol.* 31: 49–56, 1988.

STEVENS, S. S. AND NEWMAN, E. B. The localization of actual sources of sound. *Am. J. Psychol.* 48: 297–306, 1936.

SUGA, N. Cortical computational maps for auditory imaging. *Neural Networks* 3: 3–21, 1990.

VICTOR, J. D. AND PURPURA, K. P. Nature and precision of temporal coding in visual cortex: a metric-space analysis. *J. Neurophysiol.* 76: 1310–1326, 1996.

WENZEL, E. M., ARRUDA, M., KISTLER, D. J., AND WIGHTMAN, F. L. Localization using nonindividualized head-related transfer functions. *J. Acoust. Soc. Am.* 94: 111–123, 1993.

WORTIS, S. B. AND PFEIFFER, A. Z. Unilateral auditory-spatial agnosia. *J. Nerv. Ment. Dis.* 108: 181–186, 1948.

XU, L., FURUKAWA, S., AND MIDDLEBROOKS, J. C. Sensitivity to sound-source elevation in nontonotopic auditory cortex. *J. Neurophysiol.* 80: 882–894, 1998.

ZHOU, B., GREEN, D. M., AND MIDDLEBROOKS, J. C. Characterization of external ear impulse responses using Golay codes. *J. Acoust. Soc. Am.* 92: 1169–1171, 1992.

ZWISLOCKI, J. AND FELDMAN, R. S. Just noticeable differences in dichotic phase. *J. Acoust. Soc. Am.* 28: 860–864, 1956.