

# Auditory Cortical Images of Tones and Noise Bands

JULIE G. ARENBERG,<sup>1,2</sup> SHIGETO FURUKAWA,<sup>1</sup> AND JOHN C. MIDDLEBROOKS<sup>1,2</sup>

<sup>1</sup>*Kresge Hearing Research Institute, Department of Otorhinolaryngology, University of Michigan, Ann Arbor, MI 48109-0506, USA*

<sup>2</sup>*Kresge Hearing Research Institute, Neuroscience Graduate Program, University of Michigan, Ann Arbor, MI 48109-0506, USA*

Received: 20 March 2000; Accepted: 14 June 2000; Online publication: 29 August 2000

## ABSTRACT

We examined the representation of stimulus center frequencies by the distribution of cortical activity. Recordings were made from the primary auditory cortex (area A1) of ketamine-anesthetized guinea pigs. Cortical images of tones and noise bands were visualized as the simultaneously recorded spike activity of neurons at 16 sites along the tonotopic gradient of cortical frequency representation. The cortical image of a pure tone showed a restricted focus of activity along the tonotopic gradient. As the stimulus frequency was increased, the location of the activation focus shifted from rostral to caudal. When cochlear activation was broadened by increasing the stimulus level or bandwidth, the cortical image broadened. An artificial neural network algorithm was used to quantify the accuracy of center-frequency representation by small populations of cortical neurons. The artificial neural network identified stimulus center frequency based on single-trial spike counts at as few as ten sites. The performance of the artificial neural network under various conditions of stimulus level and bandwidth suggests that the accuracy of representation of center frequency is largely insensitive to changes in the width of cortical images.

**Keywords:** auditory cortex, guinea pig, tonotopy, neural ensembles, functional imaging

## INTRODUCTION

Neurons in the primary auditory cortex are sharply tuned for tone frequency, and the best frequencies of neurons vary systematically across the cortex. This “tonotopic” organization has been described in many species including human (Lauter et al. 1985; Pantev et al. 1994), cat (Merzenich et al. 1975; Schreiner et al. 1992), and guinea pig (Hellweg et al. 1977; Redies et al. 1989). Based on the response properties of individual neurons, one might infer that a tonal stimulus would elicit a restricted focus of cortical activity that would vary in cortical location according to the frequency of the tone. That is, the location of maximum cortical activity would signal the center frequency of a stimulus. At least two factors complicate this simple inference. First, as the stimulus level is increased, the excitatory frequency bandwidth of most cortical neurons increases (Sutter and Schreiner 1992; Redies et al. 1989). Therefore, one would expect sounds at high levels to activate larger populations of neurons than sounds at low levels. For that reason, it is not obvious that the frequency of a high-level tone could be identified from the location of an activated cortical area. Second, under some conditions cortical neurons show nonmonotonic spike-count-versus-level functions. For that reason, changes in sound level might lead to shifts in the locus of maximum activity. Indeed, Phillips et al. (1994) found in the cat auditory cortex that the loci of maximum cortical activity varied in an irregular manner as the sound pressure level (SPL) of the tone was varied. The variation in loci of maximum activity was most conspicuous in the cortical isofrequency dimension (i.e., perpendicular to the tonotopic axis) but questions remain concerning a possible impact of stimulus SPL on frequency representation in the tonotopic domain.

---

*Correspondence to:* John C. Middlebrooks • Kresge Hearing Research Institute • 1301 E. Ann Street • Ann Arbor, MI 48109-0506. Telephone: (734) 763-7965; fax: (734) 764-0014; email: jmidd@caumich.edu

The distribution of tone-evoked activity in the guinea pig auditory cortex has been mapped previously using functional optical imaging techniques. Results obtained by the imaging of voltage-sensitive dyes by two groups of investigators were somewhat inconsistent with expectations based on direct measurements of neural spike activity. Taniguchi and colleagues (Taniguchi et al. 1992; Taniguchi and Nasu 1993) described rather diffuse regions of activated cortex that shifted in location on a scale of about 1 mm over  $> 10$  ms of poststimulus time. Those investigators concluded that the tonotopic response that they observed was "transient." They also reported that increases in stimulus levels resulted in a rostral expansion of the area of activated cortex, i.e., toward neurons with lower characteristic frequencies. In contrast, studies of the tuning curves of cortical neurons would have predicted a caudal expansion. Similar results regarding time-dependent changes in the distribution of active neurons and level sensitivity were reported by Uno et al. (1993). Bakin et al. (1996) used optical imaging of an intrinsic signal through a thinned skull. Their results are more like what one would predict from neural recordings. Nevertheless, that study did not address the influence of stimulus level on activation patterns, nor did the temporal resolution of the recording method permit any conclusion regarding changes in activation patterns in poststimulus time.

The goal of the present study was to test the hypothesis that stimulus frequencies are coded by the distribution of cortical activity along the tonotopic axis. Existing accounts of tonotopic organization of single units are consistent with that hypothesis, but optical imaging studies fail to provide strong support for the hypothesis, and a previous study that examined the cortical distribution of activated single units (Phillips et al. 1994) seems to refute it. The issue of cortical coding of the place of cochlear activation also bears on the understanding of cortical responses to electrical cochlear stimulation with a cochlear prosthesis. It is generally assumed that electrical stimuli that provide broader cochlear activation produce broader cortical activation than stimuli that produce more focal cochlear activation. Broader cortical activation, in turn, might be assumed to result in less accurate representation of place of cochlear stimulation.

We tested the impact of breadth of cortical activation on the representation of place of cochlear stimulation by examining responses to acoustic stimuli that varied in level and bandwidth. We employed silicon-substrate thin-film recording probes that permitted simultaneous recording of spike activity of neurons at multiple cortical sites (Drake et al. 1988; Najafi et al. 1985). In the present study, 16 recording sites were spaced at 100- $\mu$ m intervals along the tonotopic gradient. The stimulus-specific distribution of neuronal

spikes in cortical place and poststimulus time is referred to here as the *cortical image* of a given stimulus. We measured cortical images of sounds of various pressure levels and of bandwidths varying from pure tones to broadband noise. A procedure was developed to quantify the accuracy of stimulus representation by cortical images sampled at 16 sites. The accuracy of identification of stimulus center frequencies provided an empirical measure of the accuracy of cortical frequency representation.

Our results obtained with pure tone stimuli confirmed the inference that the cortical image of a pure tone consists of a restricted focus of activity that shifts systematically as a function of stimulus frequency. When the extent of the cochlear activation was broadened either by increasing sound level or by broadening the stimulus bandwidth, the extent of the cortical image increased. Cortical images sampled at no more than 16 sites signaled the stimulus center frequency with reasonable accuracy. The accuracy of center-frequency identification was invariant with changes in stimulus levels and bandwidths, which demonstrates the robustness of frequency-related information carried by small ensembles of cortical neurons.

## METHODS

### Anesthesia and surgery

Data were collected from 13 healthy adult pigmented guinea pigs (500–900 g). Animals were initially sedated with a subcutaneous injection of a 3:2 mixture of ketamine hydrochloride (100 mg/ml) and xylazine (100 mg/ml). Additional intramuscular injections of a 4:1 mixture of ketamine/xylazine were given as needed to maintain an areflexive state. Core body temperature was maintained at 38°C with a thermostatically controlled heating pad. A tracheal cannula was inserted. A head holder was mounted anterior to bregma. The temporalis muscle was reflected, the skull was opened on the right side, and a small hole was made in the dura over the primary auditory cortex [area A1, sometimes referred to as area A by other authors (e.g., Redies et al. 1989)]. All procedures were in accordance with the policies of the University of Michigan's University Committee on Use and Care of Animals.

### Stimulus generation and calibration

Experiments were controlled by an Intel-based personal computer. Acoustic stimuli were synthesized digitally using equipment from Tucker-Davis Technologies (TDT) (Gainesville, Florida). The sample rate for audio

output was 100 kHz with 16-bit resolution. Experiments were conducted in a sound-attenuating chamber. Sound stimuli were presented monaurally to the ear contralateral to the studied cortical hemisphere. A headphone enclosed in a small case was connected to a sound-delivery tube placed in the external auditory meatus near the tympanic membrane. The headphone was calibrated using a 1/8-in. Bruel & Kjaer condenser microphone (Naerum, Denmark) and a 0.3-cc coupler. The resulting calibration table was used for online correction of the headphone response. Sound bursts were 80–100 ms in duration. Tones were ramped on and off with 5-ms rise/fall times, and noise bursts had 0.5-ms rise/fall times. Narrowband noise stimuli were Gaussian noises filtered with trapezoid amplitude spectra having 3-dB bandwidths of 1/6, 1/3, and 1 octave and skirts that fell off at 100 dB/octave. Center frequencies of noise bands are given as the geometric means of low- and high-frequency cutoffs. Broadband Gaussian noise bursts had passbands of 1–30 kHz with abrupt cutoffs. Sound levels of tones and noise bands were equated for root-mean-squared power.

One animal was used for a survey of tonotopic organization with three probe placements (i.e., 48 recording sites). Twelve animals with a total of 15 probe placements were used for detailed study of cortical images. The sets of center frequencies, bandwidths, and levels varied from animal to animal. Stimulus center frequencies ranged from as wide as 1–30 kHz in either 1/3- or 1/6-octave steps. Levels ranged from 0 to as high as 90 dB sound pressure level (SPL) in 10-dB steps.

### Multichannel recording and spike sorting

We used silicon-substrate, thin-film, multichannel recording probes to record unit activity (Center for Neural Communication Technology, Ann Arbor, MI) (Drake et al. 1988; Najafi et al. 1985). Each probe had 16 recording sites along a single shank at 100- $\mu\text{m}$  intervals (center to center). The shank was 15  $\mu\text{m}$  thick and 100  $\mu\text{m}$  wide, tapering in width from 100 to 15  $\mu\text{m}$  over the segment containing the recording sites. The multichannel probe permitted simultaneous recording of spike activity from all sites. The impedances at each site were 1.5–4  $M\omega$ . In the guinea pig area A1, neurons sensitive to high frequencies are situated dorsocaudally, and low-frequency neurons are situated ventrorostrally (Hellweg et al. 1977; Redies et al. 1989). The recording probe penetrated the cortex from dorsocaudal to ventrorostral, roughly parallel to the cortical surface, with the width of the shank roughly parallel to the radial cell columns (i.e., with the wide axis of the probe perpendicular to the cortical surface). We attempted to position the probe in the cortical layers

that are most active under anesthesia, presumably layers III and IV, and parallel to the tonotopic gradient along which the frequency tuning of neurons changes most rapidly. Prior to detailed study at any recording-probe placement, tuning properties of rostral, middle, and caudal sites were estimated to verify proper probe placement relative to the tonotopic map in area A1. If the reverse tonotopic order was detected, indicative of the dorsocaudal field (area DC), the probe was retracted and placed further rostral in area A1.

Signals from the recording probe were amplified with a custom 16-channel amplifier, digitized at a 25-kHz rate, sharply low-pass filtered below 6 kHz, resampled at a 12.5-kHz sample rate, and then stored on the computer hard disk. Unit activity was isolated from the digitized signal off-line using custom spike-sorting software (Furukawa et al. 2000). Spike times were stored at 20- $\mu\text{s}$  resolution for further analysis. We sometimes encountered well-isolated single units, but most recordings were of small clusters of unresolved units. Recordings at particular sites were excluded from further analysis if units did not respond to any stimulus with an average of at least 1 spike/trial or if the mean spike count changed by more than a factor of 2 over the recording period.

### Data analysis

Spike counts were normalized at each recording site according to the 95th percentile of the average spike counts computed within one stimulus type (i.e., across all the tones or across all the noise bands of a certain bandwidth). By normalizing in this manner, we emphasized stimulus-driven changes in activity rather than differences in absolute spike numbers across channels. The results include data from individual recording sites as well as the distribution of activity across multiple recording sites. For individual recording sites the lowest level that elicited a stimulus-locked response defined the threshold. The *best frequency* (BF) was defined as the frequency that gave the strongest response at 10 dB above the threshold. In some cases off-line analysis revealed responses at the lowest level that was tested. In those cases, threshold was noted as the lowest tested level. The *tuning bandwidth* at any particular sound level was taken as the width of the spike-count-versus-frequency function at the interpolated half-maximal spike count for that stimulus sound level. In cases in which the spike-count-versus-frequency function did not cross the half-maximal spike count at either the high end or the low end of the tested frequency range, the 50% frequency cutoff was estimated to be halfway between the 75% cutoff frequency and the next extrapolated frequency, usually 1/3 octave beyond the tested range. If the spike-count-versus-frequency function did not cross at least the

75% cutoff frequency on both sides, then the tuning bandwidth was not computed for that site. The tuning bandwidth was calculated for pure tones and for 1/6- and 1/3-octave-wide noise bands at levels 10, 20, and 30 dB above each unit's threshold.

The spatiotemporal distribution of activity across all recording sites was referred to as a *cortical image*. The cortical image of any particular sound was derived from simultaneous recordings at 16 cortical sites, so the activity across all sites reflected the response to the same stimuli. The threshold of a cortical image for a particular tone or noise band was defined as the sound level that elicited a stimulus-locked response from the most sensitive unit recorded across the 16 recording sites. The *centroid* of the cortical image was defined as the spike-count-weighted center of mass calculated for the sites at which the firing rate was at least half the maximum normalized spike count of the distribution.

We used an artificial neural network algorithm to recognize cortical images and thereby to identify the stimulus center frequencies. This analysis was similar to a previous analysis of coding of sound-source location in the cat auditory cortex (Furukawa et al. 2000). A 2-layer, feed-forward, fully connected network was implemented with the MATLAB Neural Network Toolbox (The Mathworks, Natick, MA). The network input was a vector of the spike counts on each of the 16 recording channels; spike counts were computed over a range of 10–50 ms after stimulus onset. The input vector to the neural network was the profile of spike counts across all recording sites during a single trial. The network output was an estimate of the stimulus frequency. A layer of eight hidden units had hyperbolic-tangent transfer functions and the single output unit had a linear transfer function. Supervised training of the network used the back-propagation algorithm (Rumelhart et al. 1986).

For the purpose of testing the neural network recognition of spike patterns, we sorted the spike profiles for odd- and even-numbered trials into training and test sets, respectively. Thus, 40 trials yielded 20 training trials and 20 test trials for each stimulus. The separation of training and test sets provided a cross validation of the pattern recognition scheme. Note that each spike count vector is derived from a single trial (except when specified otherwise) rather than from an average of spike counts across trials as used in some of our previous studies (e.g., Middlebrooks et al. 1998). Changes in stimulus level often produced large changes in spike patterns, typically increasing spike counts with increasing sound level. We wished to examine the effects of different stimulus levels and to identify features of cortical images that were invariant with changing stimulus intensity. For that reason, network analyses were performed separately for responses to stimuli at levels of 10, 20, and 30 dB above threshold

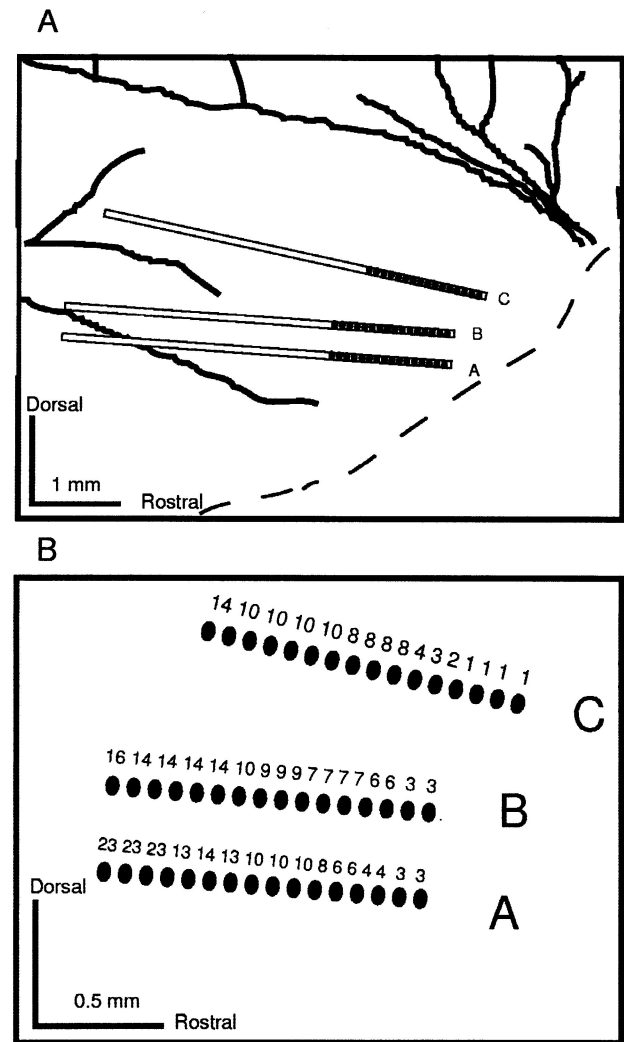


FIG. 1. Tonotopic organization of the auditory cortex. **A**. Drawing of the lateral view of the guinea pig auditory cortex. Solid lines represent surface vasculature; dashed line shows the edge of the cortical exposure. The blood vessels drawn in the dorsal-rostral corner show the location of the Sylvian fissure. The trajectories of three recording probe penetrations were reconstructed from photographs made during the physiological recording session performed in this animal. The locations of the recording sites are shown as black symbols at the rostral ends of the probes. **B**. Each of these recording sites is labeled with the best frequency (kHz) obtained at that site. Data are from GP200017.

and also for responses to stimuli that roved among those three levels.

## RESULTS

All animals showed tonotopic organization that was described previously (Hellweg et al. 1977; Redies et al. 1989). Figure 1 shows a tonotopic map that was derived from using three 16-channel probe placements. The dorsocaudal to ventrorostral orientation of

probes shown in this case was similar to the orientation used in all other animals. The numbered symbols drawn on the cortical surface represent the best frequencies of underlying recording sites in the middle cortical layers. A systematic high-to-low shift in best frequencies was observed along the caudal-to-rostral extent of each array of recording sites. Best frequencies were relatively constant in the “isofrequency” dimension perpendicular to the orientation of the probes.

We obtained detailed measurements of cortical images from 15 probe placements in 12 animals. Stable recordings of small clusters of units were obtained at 9 to 16 sites from each probe placement, a total of 202 sites. The results are presented in three subsections. First we review the sensitivity of individual sites to sounds that varied in frequency, level, and bandwidth. Next we characterize the distribution of spike activity across cortical place and poststimulus time (*cortical images of sound*). Finally, we evaluate the accuracy with which the center frequencies of tones and noise bands could be identified by their cortical images.

#### Responses of units to tones and noise bands

Units at single recording sites were selective for tone frequencies. Figure 2 represents the spike-count-frequency profiles obtained from three recording sites at one probe placement. Each panel shows the responses obtained at one site. Site locations are measured relative to the most caudal site. Solid and dashed lines represent the spike counts at levels 10 and 30 dB above threshold, respectively. The peak of the spike-count-versus-frequency profile for a stimulus level 10 dB above each unit's threshold was defined as the best frequency (BF), indicated in the figure by an arrowhead. The BFs of sites decreased monotonically from highest frequencies at the most caudal recording sites to lowest frequencies at the rostral sites. In this example, BFs of 18.9, 11.9, and 1.9 kHz were measured at 100, 800, and 1400  $\mu\text{m}$  respectively. Typically, BFs measured along the full 1500- $\mu\text{m}$  length of recording probes spanned 2–3 octaves.

The breadth of frequency tuning was quantified by the width of the spike-count-versus-frequency profiles at half the maximum spike rate. For the 92 recording sites for which tuning bandwidths were measured, the bandwidth at 10 dB above threshold was  $0.96 \pm 0.41$  (mean  $\pm$  S.D.) octaves, increasing to  $1.47 \pm 0.63$  octaves at 30 dB above threshold.

In six of the guinea pigs (79 recording sites), we also measured responses to bandpass noise at various center frequencies, bandwidths, and sound levels. Bandwidths were 1/6, 1/3, 1 octave, and broadband. Absolute thresholds for individual sites across bandwidths had a range of  $\leq 10$  dB for 67 of 79 sites and a range of 20–30 dB at the remaining 12 sites. On average,

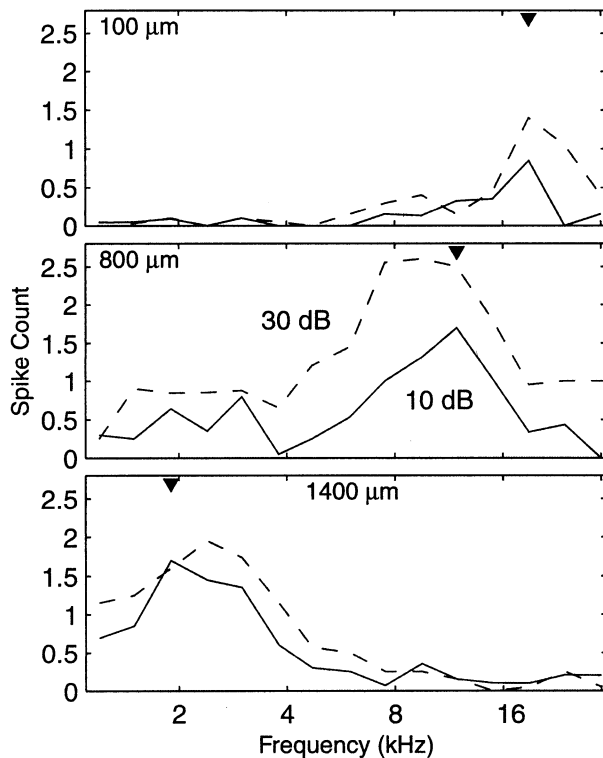


FIG. 2. Spike count versus frequency profiles. Each panel represents data from one recording site. Site locations are recorded in microns relative to the most caudal site. Solid and dashed lines represent stimulus levels of 10 and 30 dB above threshold, respectively. Triangles represent best frequencies for 10-dB stimulation. Data are from animal 9909.

sites responded more strongly to all noise bandwidths than to tones under conditions in which sound levels were equalized for root-mean-squared power and in which tone frequencies and noise-band center frequencies were adjusted to optimum values. At levels 20 dB above the BF-tone thresholds, spike counts (not normalized) elicited by narrowband and broadband noises were significantly greater than the spike counts elicited by BF tones. Specifically, the ratio of noise-band and spike counts to BF-tone spike counts was  $137 \pm 71$ ,  $141 \pm 83$ ,  $129 \pm 63\%$  (mean  $\pm$  S.D.) for 1/6-, 1/3-, and 1-octave noise bands, respectively ( $p < 0.01$ , paired *t*-test) and  $120 \pm 70\%$  for broadband noises ( $p < 0.05$ ). Plots of spike count as a function of noise-band center frequency showed broader frequency tuning than for tones. Tuning bandwidths for stimuli 20 dB above threshold were  $0.99 \pm 0.41$  octaves for tones,  $1.91 \pm 0.68$  octaves for 1/6-octave noise bands, and  $1.95 \pm 0.64$  octaves for 1/3-octave noise bands ( $N = 46$  sites). Tuning bandwidths were not calculated for 1-octave noise bands because the tuning often extended beyond the frequency range that was tested, resulting in an inadequate number of sites for this analysis.

Most spike counts increased monotonically with

increasing sound level. Responses were considered nonmonotonic if responses to the maximum level tested were less than 75% of the maximum response to any lower sound level. Three animals with five probe placements were tested with a range of pure-tone stimulus levels extending up to 50–90 dB above threshold. In response to BF tones, only 2 of 53 recording sites were nonmonotonic. Off-best-frequency stimuli presented at 50–90 dB above threshold resulted in nonmonotonic rate level functions for 9 of 200 recording-site–frequency combinations. In two animals in which two parallel probe placements were tested, there were no differences in the proportion of monotonic versus nonmonotonic units across placements. For nine animals, complete data were obtained for tones and noise bands up to only 30 dB above the threshold for each tone or noise condition. For BF tonal stimuli, only 7% of sites showed responses at 30 dB above threshold that were less than 75% of the maximum response to any lower sound level. Thus, across all cases studied, a very small percentage of sites demonstrated nonmonotonic spike-count-level functions.

#### Cortical images of tones and noise bands

We refer to the characteristic spatiotemporal distribution of cortical spike activity elicited by any particular stimulus as the “cortical image” of that stimulus. For each stimulus frequency, threshold was defined as the stimulus level that elicited stimulus-locked spikes from the most sensitive recording site. Figure 3 presents cortical images of pure tones at three frequencies at levels 20 dB above threshold. In each panel, the activity at each recording site was averaged across the 40 presentations of a particular tone frequency. The vertical axis represents the location along the tonotopic axis relative to the most caudal recording site, the horizontal axis represents time after stimulus onset, and colors represent the normalized spike probability. In the examples shown in Figure 3, the cortical images of tones each contained a single focus of activity at which sites responded with highest probability and shortest latency. Spike probabilities decreased and latencies increased gradually with increasing distance rostrally and caudally from the focus. The dispersion of first-spike latencies across the cortical image was  $\leq 5$  ms. The focus of activation shifted in the cortex from rostral to caudal as the stimulus frequency was shifted from low to high.

The cortical images of tones increased in width in response to increasing sound levels. Figure 4 presents cortical images of a 10.1-kHz tone at four stimulus levels relative to threshold. As the sound level was increased, the peak activity in the cortical image increased monotonically. The cortical image broadened further caudally than rostrally leading to a slight

caudal shift in the overall distribution of activity. The asymmetric broadening of the image indicates greater recruitment of sites with BFs higher than the stimulus frequency.

As the stimulus bandwidth was increased from that of a pure tone to 1 octave, the width of the activated cortical area increased. Figure 5 shows an example of the growth of the cortical image with increasing stimulus bandwidth. Each panel represents the image of a tone or noise band centered at 2 kHz presented at 20 dB relative to threshold for each stimulus. The width of the activated cortical image area increased from a few sites to encompass nearly all of the recorded sites.

Cortical image widths were quantified by computing the area under the curve of normalized spike count versus cortical place, then dividing the area by the maximum normalized spike count. The result is the width of a rectangle of unit height with area equal to the area under the spike-count curve. Figure 6 represents the image width versus stimulus level for one animal averaged across four frequencies. Image widths of tones and noise bands increased considerably as the stimulus level was increased as well as when the bandwidth was increased from that of a tone to 1 octave. The widths of images of 1-octave bands are somewhat underestimated, particularly at high stimulus levels, because the cortical activation often extended beyond the range of the recording probe for these stimuli. The size of the cortical image of a sound of particular center frequency, level, and bandwidth varied considerably among five animals. We presume that such variation reflected individual differences in the details of the tonotopic organization and differences in the precise orientation of our recording probes relative to the tonotopic axis and to cortical layers. Nevertheless, the trends of larger cortical image widths for increased sound levels and bandwidths could be generalized for all animals.

The focus of maximum activity varied in location according to the center frequency of the stimulus. We quantified the location of the focus of activity by the centroid. Figure 7 represents the centroid as a function of the stimulus frequency in one animal. Each symbol represents a particular sound level. For low stimulus levels, the centroid shifted roughly linearly from rostral to caudal as the stimulus frequency was changed from low to high on a logarithmic scale. As the stimulus level was increased, centroids of images of lower-frequency tones tended to shift caudally. That can be seen in the example illustrated in Figure 4. In Figure 7, centroids were not computed for some of the frequencies and highest levels because the cortical images widened to extend past the range of the recording array. In each of three animals in which two parallel probe placements were made, the progression of

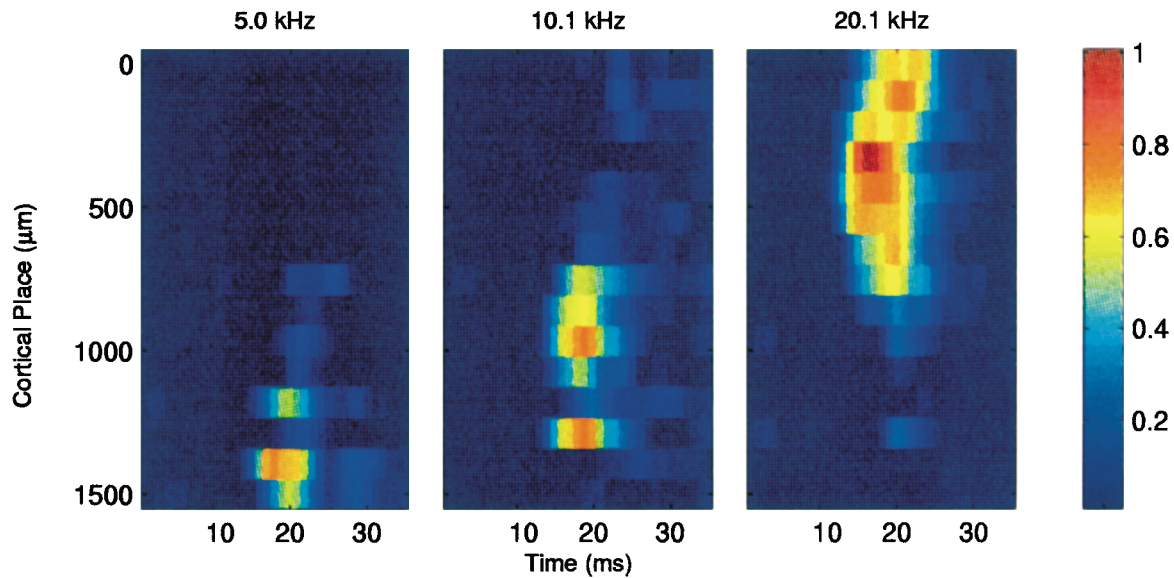


FIG. 3. Cortical images of pure tones at three frequencies. Each panel represents the cortical image of a tone at one frequency at a level of 20 dB above threshold. For each cortical image, the abscissa represents poststimulus time and the ordinate represents cortical place relative to the most caudal recording site. Normalized spike probabilities are represented by colors. Data are from animal 9926.

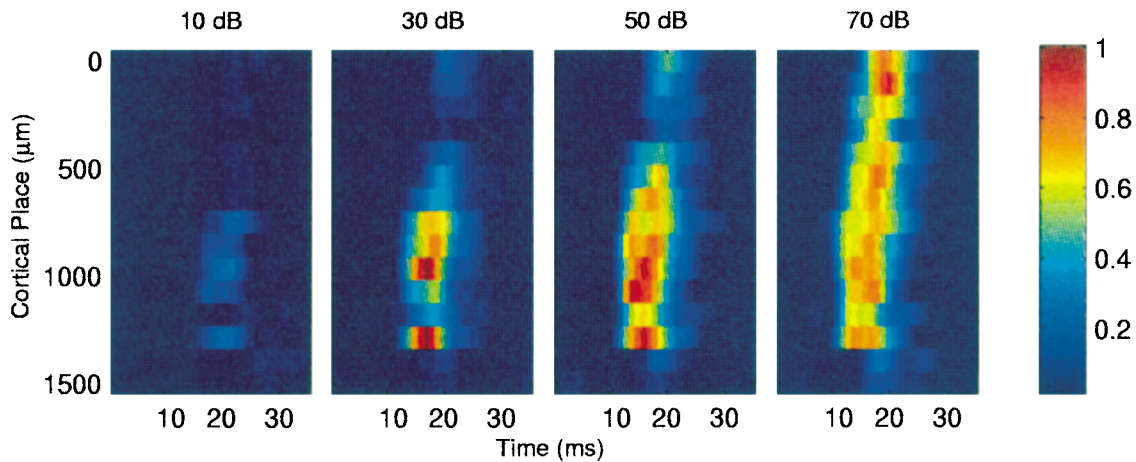


FIG. 4. Cortical images of a 10.1-kHz tone at four stimulus levels relative to threshold. Conventions as in Fig. 2. Data are from animal 9926.

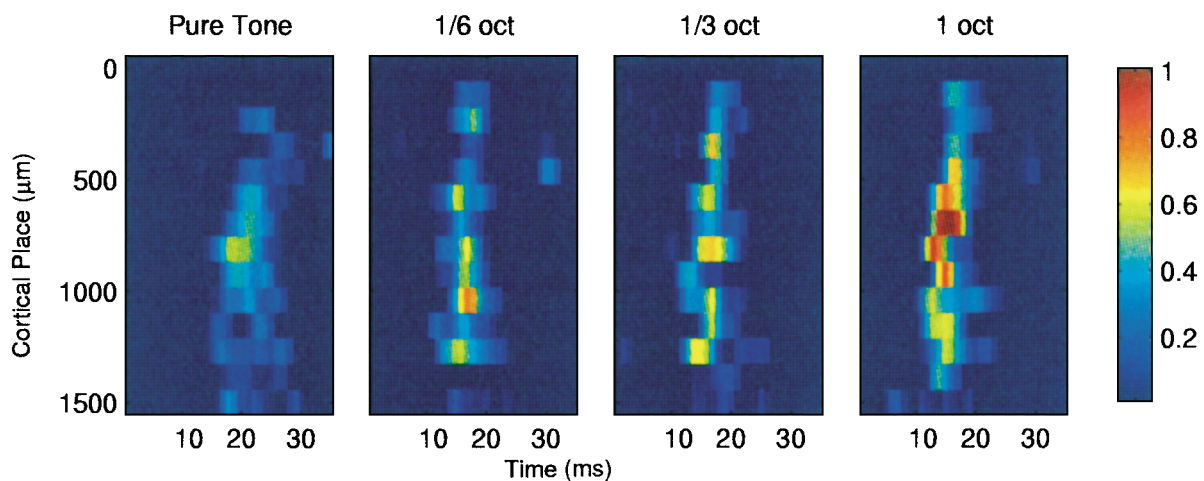


FIG. 5. Cortical images of tones and bandpassed noises centered at 2 kHz. Sound levels were 20 dB above threshold. Conventions as in Fig. 2. Data are from animal 9803.

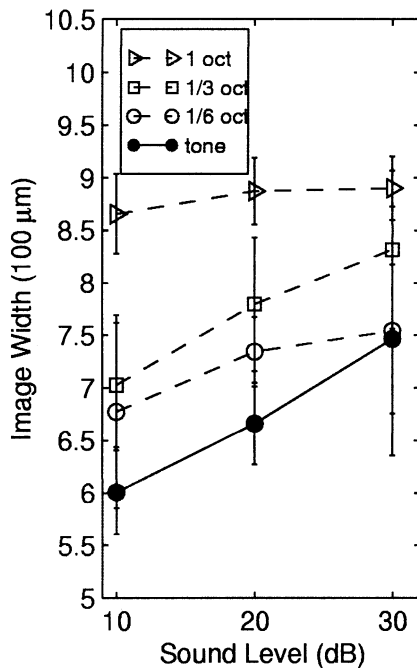


FIG. 6. Widths of cortical images as a function of sound level. Line types represent tone or noise bandwidths as indicated in the key. Each symbol represents an average of the image width for four frequencies. Error bars represent the standard deviation. Data are from animal 9803.

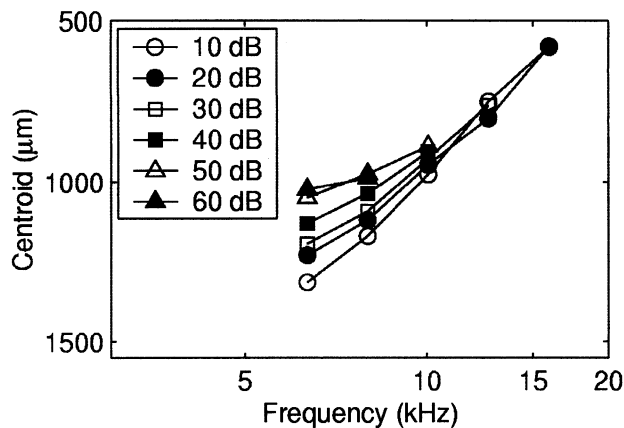


FIG. 7. Centroid of cortical images as a function of tone frequency. Each symbol represents one sound level. Data are from animal 9927.

centroid as a function of stimulus frequency was comparable between the two placements. Figure 8 plots the centroid location versus stimulus center frequency for sounds of various bandwidths. Each symbol represents a different stimulus bandwidth, and all stimuli were presented at 20 dB relative to threshold. Despite the prominent growth of cortical image width in response to increases in stimulus bandwidth, there was a relatively small effect on the frequency sensitivity of centroid locations.

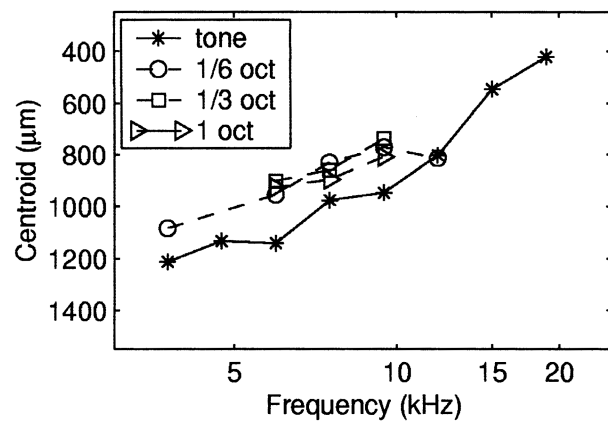


FIG. 8. Centroid of a cortical image as a function of center frequency for tones and bandpassed noises. Sound levels were 20 dB above threshold. Data are from animal 9909.

We computed the cortical scale factor between cortical place and stimulus frequency by calculating the slope of the plot of centroid location versus tone frequency ( $\mu\text{m}/\text{oct}$ ) across 11 probe placements in ten animals. For pure tones presented at 20 dB above threshold, the median scale factor was  $345 \mu\text{m}/\text{oct}$  ranging from  $203$  to  $712 \mu\text{m}/\text{oct}$ . The median scale factor is close to the value of  $\sim 375 \mu\text{m}/\text{oct}$  that can be estimated from the BF data illustrated by Redies and colleagues (Redies et al. 1989, Fig. 1). The value of  $\sim 650 \mu\text{m}/\text{oct}$  that can be estimated from the optical imaging results illustrated by Bakin et al. (1996, Fig. 9) is within the range of scale factors we computed.

#### Identification of stimuli based on trial-by-trial cortical images

We tested the accuracy with which cortical images could represent stimulus frequencies on single trials. We used an artificial neural network to recognize patterns of cortical activity and thereby to identify stimulus frequencies. A similar approach has been used previously to test the coding of sound-source locations by the temporal spike patterns of single cortical neurons (Middlebrooks et al. 1994, 1998) and by small ensembles of cortical neurons (Furukawa et al. 2000) and to test coding of locations of tactile stimuli (Nicolis et al. 1998). As in our previous work, we employed a cross-validation procedure in which cortical responses obtained on odd-numbered trials were used to train the network and the trained network was used to classify responses obtained on even-numbered trials.

We compared several experimental conditions and a control condition. In the *single-level* conditions, the stimulus levels were held constant at 10, 20, or 30 dB relative to threshold. In the *roving-level* condition, the data set contained responses to stimulus levels that varied across 10, 20, and 30 dB above threshold. In



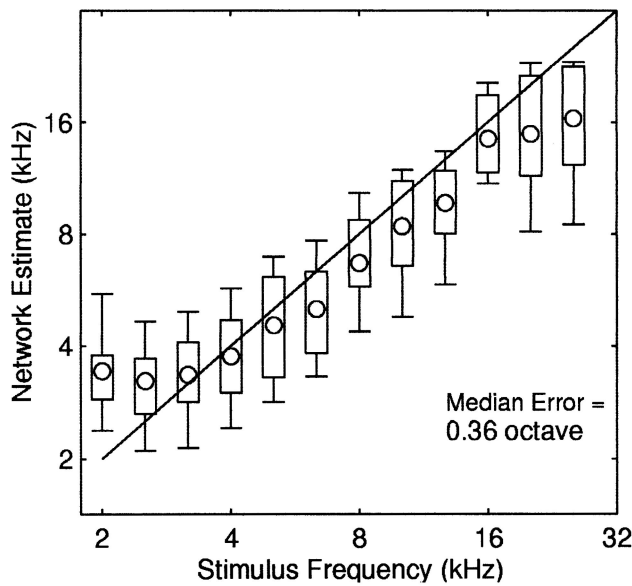


FIG. 9. Example of artificial neural network recognition performance. The abscissa represents the stimulus frequency, and the position of each box represents the distribution of network estimates of the tone frequency. Open circles indicate the mean. The boxes represent the 50% confidence intervals of the means, and the error bars represent the 90% confidence intervals. The solid diagonal line represents perfect performance. Data are from animal 9909.

the *chance performance* control condition, the network was trained and tested with spike-count profiles that were randomly reassigned to stimuli. Chance performance levels varied among animals because of inter-animal differences in the number of sites that were recorded and in the ranges of frequencies that were tested. The frequency range tested for each animal was limited to the range of BFs encountered in that probe placement. The network was trained and tested separately for ten probe placements in nine animals. For each probe placement, the same range of stimulus frequencies was tested in all level conditions. A higher threshold for at least one stimulus frequency limited some probe placements to two single-level conditions of 10 and 20 dB and limited the roving-level condition to a range of 10–20 dB across trials.

Figure 9 shows an example of identification of tone frequencies in the roving-level condition based on 16-channel recordings in one probe placement. The abscissa shows the actual tone frequencies, and the position of the boxes along the ordinate represents the distribution of network estimates of the stimulus frequency. The median of the unsigned error across all stimulus frequencies (the *median error*) in this case was 0.36 octave, which was less than half of the 0.99-octave median error obtained in the chance-performance condition.

When results of this analysis were compared across animals, the network performance was relatively stable

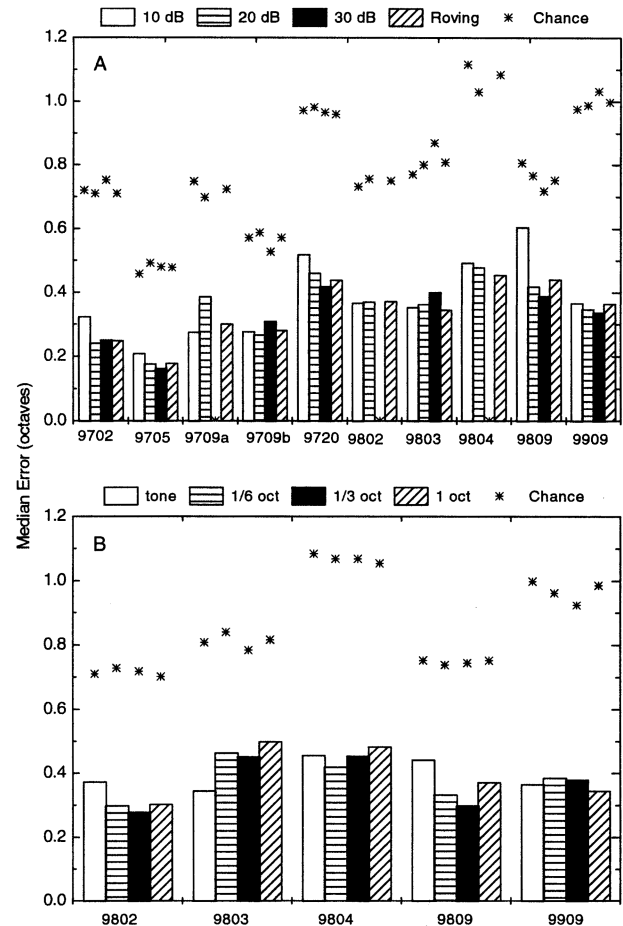


FIG. 10. Comparison of network performance across animals and stimulus types. Each cluster of bars represents data for a single animal, which is represented by a number. Lower-case a and b indicate two probe placements in one animal. Plotted are the median errors in octaves for (A) single-level and roving-level conditions and for (B) tones and bandpassed noises. The asterisk above each bar indicates the chance performance level for that condition.

across various fixed and roving stimulus levels. The bar plot in Figure 10A represents the median error of network performance for ten probe placements in nine animals for pure tones only. Each cluster of bars represents the measurements for one placement for the 10-, 20-, and 30-dB single-level condition and for the roving-level condition. The asterisks represent the median error of the chance performance for each stimulus condition. The network performance varied for different animals, in part because of a different number of frequencies and levels included in the analysis. For all level conditions across animals, the network median error ranged from 0.16 to 0.6 octave and generally was about half the chance error levels. The accuracy of frequency identification showed no systematic dependence on stimulus level or single-versus roving-level conditions.

In five of the animals for which narrowband noise

stimuli were presented, we also examined the artificial-neural-network identification of stimulus center frequency for tones and narrowband noises in the roving-level condition (Fig. 10B). Each cluster of bars represents the measurements for one animal, and each bar in a cluster represents an individual bandwidth condition. For all bandwidths, the network performance ranged from 0.28 to 0.5 octave in median error. Again, the network median errors were about half the chance levels. The accuracy of center-frequency identification and the variance of that measure showed no systematic dependence on stimulus bandwidth. That is, as was observed for increased stimulus level, increases in the widths of cortical images resulting from increased bandwidths did not alter the network performance.

The accuracy of stimulus representation by spike count presumably was limited by trial-by-trial variability in responses. We attempted to obtain more precise estimates of spike probabilities by averaging across trials. Given a training set (or a test set) of 20 responses to each stimulus condition, we formed a bootstrap average of each cortical image by repeatedly drawing  $n$  samples with replacement from the set of cortical images elicited by stimuli of a particular frequency and level (Efron and Tibshirani 1991). Because we sampled with replacement, each bootstrap average could contain zero, one, two, or more instances of each cortical image. For each of ten recording probe placements in nine animals, we formed 20 bootstrapped training patterns and 20 bootstrapped test patterns for each stimulus condition. The accuracy of center-frequency identification by the artificial neural network was tested using across-trial averages of  $n = 1$  to 2048 trials. Median errors approached an asymptote by  $n = 16$  to 64 trials. The asymptotic levels of performance ranged from median errors of 0.09–0.22 octave. Generally, the asymptotic median errors given an averaged input were less than half the median errors obtained on a trial-by-trial basis and less than a quarter of the chance-level performance.

## DISCUSSION

The results of our studies demonstrate that (1) pure tone stimuli elicit a restricted focus of activity along the tonotopic axis of cortical area A1; (2) the location of the activation focus shifts as a function of the tone frequency; (3) the width of the activation focus broadens as the stimulus level or bandwidth is increased; (4) the stimulus center frequency can be identified with some accuracy on a trial-by-trial basis of spike counts at as few as ten recording sites; (5) stimulus identification is robust to changes in the stimulus level or bandwidth. This section relates these results to those

of previous studies, then considers properties of cortical coding of stimulus frequency by small ensembles of neurons and implications for certain hypotheses regarding cochlear prostheses.

### Relation to previous studies

Previous studies have examined the frequency responses of auditory cortical neurons in the guinea pig (Hellweg et al. 1977; Redies et al. 1989). The current study confirms the previously reported distribution of best frequencies, with low frequencies represented rostrally and high frequencies caudally. The sharpness of tuning and the response latencies were comparable to those reported previously.

Several studies have examined cortical responses to tones at moderate-to-high stimulus levels in anesthetized guinea pigs (Taniguchi et al. 1992; Taniguchi and Nasu 1993; Uno et al. 1993; Bakin et al. 1996). These studies employed optical recording techniques that offer the benefit of a two-dimensional functional image at the expense of reduced spatial and temporal resolution. Results obtained with voltage-sensitive dyes showed that the area of greatest change in optical signal varied in cortical location on a scale of roughly 1 mm over the course of greater than 10 ms of post stimulus time, i.e., that the tonotopic representation was transient (Taniguchi et al. 1992; Taniguchi and Nasu 1993; Uno et al. 1993). Our recordings of neural spike activity failed to replicate that result. In our results, spike activity typically was restricted to a burst lasting no more than about 10 ms and the dispersion of first-spike latencies across cortical recording sites was no more than 5 ms. The voltage-sensitive-dye technique also showed that as sound levels increased the area of activated cortex expanded rostrally toward neurons with lower characteristic frequencies. In contrast, we observed a caudal expansion of cortical images. The result from our unit recordings is consistent with the recruitment of high-frequency neurons by tones activating the low-frequency tails of their tuning curves. A previous study in the guinea pig auditory cortex showed that the low-frequency slopes of tuning curves tend to be shallower than the high-frequency slopes (Redies et al. 1989). Our results are generally consistent with the results obtained from optical imaging of intrinsic signals (Bakin et al. 1996). Quantitative comparisons are difficult because of the differences in methods. However, the extent of activation along the tonotopic dimension of the cortex seems roughly comparable in the two studies, and the tonotopic scale factor that we estimate from one example in the Bakin study is within the range of scale factors obtained in our sample of ten animals. The Bakin study showed regions of activation that formed parallel bands elongated along a dimension orthogonal to the

tonotopic axis. Consistent with that result, we observed cortical images that were about equally wide regardless of the location of the recording probe in the iso-frequency dimension.

The previous study most similar to the present one used numerous sequential electrode penetrations to map the cortical images of tones in barbiturate-anesthetized cats (Phillips et al. 1994). In that study, many neurons showed nonmonotonic rate-level functions, and nonmonotonic neurons tended to segregate from neurons that showed monotonic rate-level functions. As a consequence of the nonmonotonic rate-level functions, the cortical images tended to change in shape and location as the stimulus level was varied at a constant frequency. In contrast, in the ketamine-anesthetized guinea pig we seldom observed nonmonotonic rate-level functions and, aside from a tendency for centroids to shift somewhat caudally at increasing sound levels, we saw little change in the loci of maximum activity as stimulus levels increased. Several differences in the experimental design might have contributed to the differences in results. Of course, the difference in species is a possible factor. Also, we used ketamine anesthesia, whereas barbiturate anesthesia was used in many of the studies that found nonmonotonic rate-level responses (Sutter and Schreiner 1991; Phillips et al. 1994; Heil et al. 1994). Finally, the cat study was subject to changes in the physiological state of the animal over sequential electrode penetrations made over the course of 47–52 hours (Phillips et al. 1994). In contrast, we recorded simultaneously from all sites and interleaved stimuli of various frequencies, levels, and bandwidths.

#### Cortical representation of cochlear place of stimulation

Cortical images based on responses averaged across trials showed clear centroids of maximum activity that shifted in cortical location according to stimulus center frequency. That is, the place of cortical activity represented the place of cochlear stimulation. We used an artificial neural network analysis of trial-by-trial cortical images to estimate the accuracy of cortical representation of stimulus center frequencies. We found that single-trial responses at as few as ten cortical sites could signal center frequency with accuracy well above chance levels of performance. Averages of spike counts across trials provided a more precise representation of spike probabilities and led to more accurate estimates of center frequency. Nevertheless, the asymptotic values of accuracy in frequency identification were appreciably less accurate than psychophysical difference limens (Shower and Biddulph 1931; Wier et al. 1977; Nelson et al. 1983). Accuracy in our results

presumably was limited by the size of the neural population that could be sampled, which presumably was many orders of magnitude smaller than the population that is used by a listener in a psychophysical task.

The widths of cortical images increased monotonically with increases in either stimulus level or stimulus bandwidth. The increases in cortical image widths probably produced two or more somewhat opposing effects on the accuracy of representation of center frequency. First, one might imagine that a given center frequency could be signaled by activity in a particular restricted population of neurons. In that view, increases in the width of cortical images would increase the uncertainty in center-frequency representation, leading to a decrease in the accuracy of frequency identification. In contrast, an increase in the number of active neurons might be thought of as an increase in the number of independent sources of information about stimulus frequency, thereby increasing the precision of stimulus representation. In that view, an increase in the size of the cortical image would predict an increase in the accuracy of frequency identification. The results showed that the accuracy of frequency identification based on recognition of cortical images using an artificial neural network was roughly equal across stimulus bandwidths up to 1 octave and across a range of sound levels. Any degradation in frequency coding brought about by the spread of the cortical image was balanced by the increase in the number of neurons carrying stimulus-related information. Our physiological results are comparable with results of psychophysical tests of frequency discrimination by human listeners (Shower and Biddulph 1931; Wier et al. 1977; Nelson et al. 1983). In one study (Wier et al. 1977), listeners showed a marked decrease in frequency difference limens as stimulus levels were increased from threshold to about 10–20 dB above threshold. With further increases in stimulus levels, listeners then showed much smaller decreases in difference limens, no more than a factor of 2 as levels were increased to as much as 80 dB above threshold.

#### CONCLUSIONS

This study has characterized the cortical images of simple sounds and has demonstrated an empirical procedure for estimating the accuracy of stimulus representation in the auditory cortex. We hope that this animal model will be applied to more complex sounds such as animal vocalizations and human speech. The model also has application for the study of electrical stimulation of the auditory system with a cochlear prosthesis. One school of thought in the cochlear prosthesis community holds that the most accurate representation of cochlear place of stimulation will

result from electrical current fields that produce the most restricted cochlear activation and, thus, the most restricted cortical activation. Our results using acoustical stimuli indicate that more restricted cortical activation patterns do not necessarily lead to more accurate representation of cochlear place of stimulation. Research has begun in our laboratory to investigate cortical responses to cochlear prosthesis stimulation (Arenberg and Middlebrooks 2000).

## ACKNOWLEDGMENTS

This work was supported NIH/NIDCD grants ROI DC04312 and T32 DC00011. We thank Zekiye Onsan and Chris Ellinger for their technical support and Bryan Pfingst and Steven Bierer for their comments on an earlier version of the manuscript. Multichannel recording probes were generously provided by the University of Michigan Center for Neural Communication Technology, which is supported by NIH/NCRRT grant P41-RR09754.

## REFERENCES

- ARENBERG JG, MIDDLEBROOKS JC. (2000) Cortical responses to multichannel cochlear implant stimulation. 23rd Midwinter Meeting of the Association of Research in Otolaryngology, St. Petersburg Beach, FL, February 2000.
- BAKIN JS, KWON MC, MASINO SA, WEINBERGER NM, FROSTIG RD. Suprathreshold auditory cortex activation visualized by intrinsic signal optical imaging. *Cerebral Cortex*. 6(2):120–30, 1996.
- DRAKE KL, WISE KD, FARRAYE J, ANDERSON DJ, BEMENT SL. Performance of planar multisite microprobes in recording extracellular single-unit intracortical activity. *IEEE Trans. Biomed. Eng.* BME-35:719–732, 1988.
- EFRON B, TIBSHIRANI R. Statistical data analysis in the computer age. *Science* 253:390–395, 1991.
- FURUKAWA S, XU L, MIDDLEBROOKS JC. Coding of sound source location by ensembles of cortical neurons. *J. Neurosci.* 20(03):1216–1228, 2000.
- HEIL P, RAJAN R, IRVINE DR. Topographic representation of tone intensity along the isofrequency axis of cat primary auditory cortex. *Hear. Res.* 76(1–2):188–202, 1994.
- HELLWEG FC, KOCK R, VOLRATH M Representation of the cochlea in the neocortex of guinea pigs. *Exp. Brain Res.* 29:467–474, 1977.
- LAUTER JL, HERSCOVITCH P, FORMBY C, RAICHLER ME. Tonotopic organization in human auditory cortex revealed by positron emission tomography. *Hear. Res.* 20:199–205, 1985.
- MERZENICH MM, KNIGHT PL, ROTH GL. Representation of cochlea within primary auditory cortex in the cat. *J. Neurophysiol.* 38(2):231–249, 1975.
- MIDDLEBROOKS JC, CLOCK AE, XU L, GREEN DM. A panoramic code for sound location by cortical neurons. *Science* 264:842–844, 1994.
- MIDDLEBROOKS JC, XU L, EDDINS AC, GREEN DM. Codes for sound-source location in nontopographic auditory cortex. *J. Neurophysiol.* 80:863–881, 1998.
- NAJAFI K, WISE KD, MOCHIZUKI T. A high-yield IC-compatible multichannel recording array. *IEEE Trans. Electron. Dev.* 32:1206–1211, 1985.
- NELSON DA, STANTON ME, FREYMAN RL. A general equation describing frequency discrimination as a function of frequency and sensation level. *J. Acoust. Soc. Am.* 73(6):2117–2122, 1983.
- NICOLELIS MAL, GHAZANFAR AA, STAMBAUGH CR, OLIVEIRA LMO, LAUBACH M, CHAPIN JK, NELSON RJ, KAAS JH. Simultaneous encoding of tactile information by three primate cortical areas. *Nat. Neurosci.* 1:621–630, 1998.
- PANTEV C, BERTRAND O, EULITX C, VERKINDT C, HAMPSON S, SCHUIERER G, ELBERT T. Specific tonotopic organizations of different areas of the human auditory cortex revealed by simultaneous magnetic and electric recordings. *Electroencephal. Clin. Neurophysiol.* 94:26–40, 1995.
- PHILLIPS DP, SEMPLE MN, CALFORD MB, KITZES LM. Level-dependent representation of stimulus frequency in cat primary auditory cortex. *Exp. Brain Res.* 102:210–226, 1994.
- REDIES H, SIEBEN U, CREUTZFELDT OD. Functional subdivisions in the auditory cortex of the guinea pig. *J. Comp. Neurol.* 282:473–488, 1989.
- RUMELHART DE, HINTON GE, WILLIAMS RJ. Learning internal representations by error propagation. *Parallel Data Processing*. M.I.T. Press, Cambridge, MA, 1986, 318–362.
- SCHREINER CE, MENDELSON JR, SUTTER ML. Functional topography of cat primary auditory cortex: representation of tone intensity. *Exp. Brain Res.* 92:105–122, 1992.
- SHOWER EG, BIDDULPH R. Differential pitch sensitivity of the ear. *J. Acoust. Soc. Am.* 3:275–287, 1931.
- SUTTER ML, SCHREINER CE. Physiology and topography of neurons with multi-peaked tuning curves in cat primary auditory cortex. *J. Neurophysiol.* 65(5):1207–1226, 1991.
- TANAGUCHI I, HORIKAWA J, MORIYAMA T, NASU M. Spatiotemporal pattern of frequency representation in the auditory cortex of guinea pigs. *Neurosci. Lett.* 146:37–40, 1992.
- TANAGUCHI I, NASU M. Spatiotemporal representation of sound intensity in the guinea pig auditory cortex observed by optical recording. *Neurosci. Lett.* 151:178–181, 1993.
- UNO H, MURAI N, FUKUNISHI K. The tonotopic representation in the auditory cortex of the guinea pig with optical recording. *Neurosci. Lett.* 150:179–182, 1993.
- WIER CC, JESTEADT W, GREEN DM. Frequency discrimination as a function of frequency and sensation level. *J. Acoust. Soc. Am.* 61:178–184, 1977.