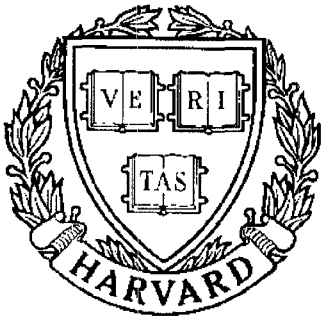
TECHNICAL RESEARCH REPORT

SYSTEMS RESEARCH CENTER



Supported by the National Science Foundation Engineering Research Center Program (NSFD CD 8803012), the University of Maryland, Harvard University, and Industry

Spectral Gradient Columns in Primary Auditory Cortex: Physiological and Psychoacoustical Correlates

by S.A. Shamma, S. Vranic and P. Wiser

TR 91-35

Spectral Gradient Columns in Primary Auditory Cortex: Physiological and Psychoacoustical Correlates

S.A. SHAMMA, S. VRANIĆ and P. WISER

Department of Electrical Engineering, Systems Research Center,
University of Maryland Institute for Advanced Computer Studies,
University of Maryland, College Park, MD 20742, U.S.A.

ABSTRACT

Mappings of the spatial distribution of responses in the primary auditory cortex (AI) reveal that both the gradient of the acoustic spectrum and sensitivity to FM sweep direction are encoded in an orderly manner along the isofrequency planes of AI. Psychoacoustical tests also demonstrate a potential perceptual correlate of the gradient maps, namely the threshold stability and heightened sensitivity of human subjects to the detection of changes in the symmetry of spectral peaks.

KEYWORDS

Spectral gradients; FM selectivity; Spectral shape; Symmetry; Bandwidth.

The shape of the acoustic spectrum is a fundamental cue in the perception and recognition of complex sounds such as speech and music. It is largely unknown, however, how this spectrum is represented in the auditory system, and what specific features are extracted and emphasized by such a representation. Existing physiological data have primarily addressed the encoding of complex sounds at relatively early levels in the auditory pathway [*Blackburn and Sachs*, 1990; *Kim et al.*, 1991; *Sachs and Young*, 1979]; while psychoacoustical studies have mostly dealt with the detection of basic spectral shapes, e.g., edges and sinusoidal ripples [*Bernstein and Green*, 1987].

The physiological experiments reported here examine the encoding of the acoustic spectrum at the level of the primary auditory cortex (AI). Specifically, we have measured the distribution of inhibitory and excitatory responses across the surface of AI. The results reveal that specific spectral features, such as the gradient of the acoustic spectrum, are extracted and mapped in AI. Based on these maps, psychoacoustical experiments were carried out to test directly the sensitivity of human subjects to changes in these features, and to study the implications of these findings to the perception of complex spectra. These physiological and psychoacoustical experiments are described in the next two sections.

The primary auditory cortex is essential for the perception and localization of sound. Our knowledge of the functional organization of AI, however, remains extremely limited. At present, only two general organizational features are firmly established: the spatially ordered tonotopic axis [*Reale and Imig*, 1980], and the alternating bands of binaural response properties that run perpendicularly to the isofrequency planes [*Imig and Adrian*, 1977] (Fig. 1). They are roughly analogous to the retinotopic maps and the ocular dominance columns of the primary visual cortex [*Hubel and Wiesel*, 1962].

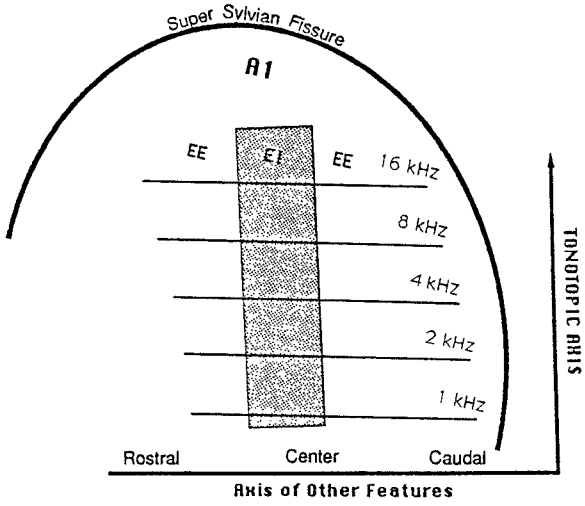


Fig. 1. Schematic of ferret primary auditory cortex (AI). The tonotopic axis runs in the mediolateral direction with low frequencies laterally. The arc represents the location of the supra-sylvian fissure. Isofrequency planes extend along the rostro-caudal axis. Presumed binaural columns intersect the isofrequency planes. Dimensions of AI vary considerably across animals, but average distance between octave frequences is 0.5-1 mm.

While these axes are fundamental to the organization of AI, they only relate to basic simple properties of the acoustic stimulus that are already established at much lower levels of the auditory pathway. Tonotopic order, for instance, originates at the cochlea, while binaural columns exist at least as early as the inferior colliculus [*Wenstrup, Ross and Pollak*, 1986]. With the exception of the more specialized auditory system of the bat [*Suga*, 1984], ordered responses to more complex stimulus features, analogous to the orientation columns and direction of motion selectivity in the visual cortex, have not been found in AI.

In a recent series of experiments, we explored this issue through a detailed mapping of the response areas of cortical cells across the surface of AI in the ferret [*Shamma, Fleshman and Wiser*, 1991]. Specifically, the aim was to establish whether any systematic changes in the balance of inhibitory and excitatory responses occur in cells along the isofrequency planes and, if so, to determine the implications of these changes to responses to frequency-modulated (FM) tones and spectrally shaped noise stimuli.

Organization of Response Areas in AI

The *response area* of a cell consists of its excitatory and inhibitory responses as a function of stimulus frequency and intensity. To obtain this measure, a two-tone stimulus is used in which the first tone, $T1$, spans a wide range of frequencies about the BF to measure the excitatory responses of the cell. The second tone, $T2$, is fixed at BF to provide a level of response against which the inhibition can be detected. The response areas thus measured can be classified into three basic types (Figs. 2(a-c)). The first exhibits symmetric (often narrow and weak) lateral inhibition (Fig. 2a). The two other types are significantly asymmetric. In one, inhibition is

strong only from frequencies above the BF (Fig. 2b); in the other, the inhibition is largely from below the BF (Fig. 2c).

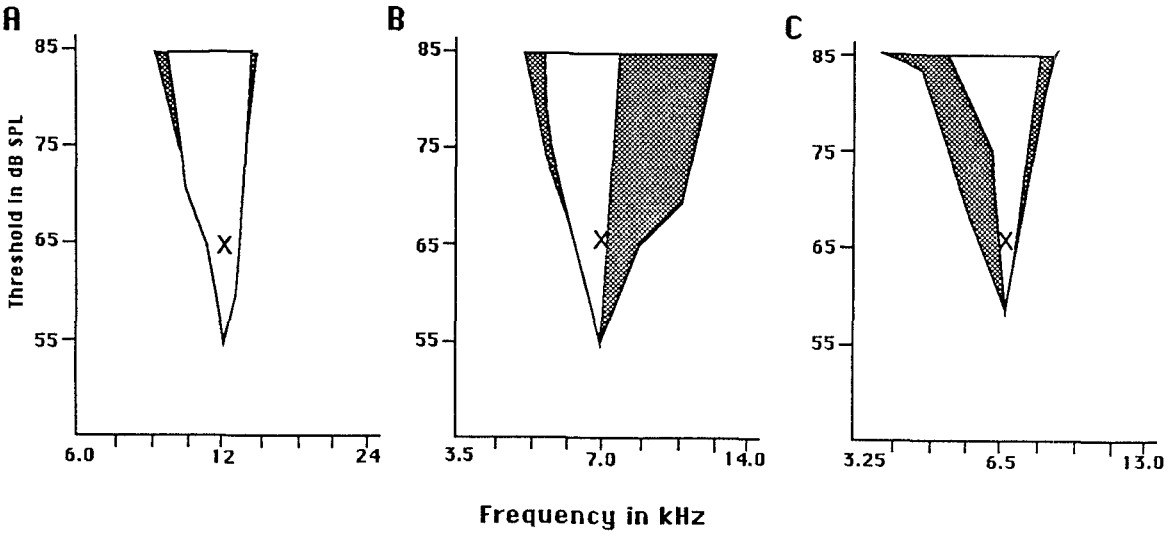


Fig. 2. Three types of response areas in AI cortical cells: (a) Response area for a cell with weak symmetric inhibition. The intensity and frequency of $T2$ in the two-tone stimulus are indicated by the "X" in the figure. (b) Response area of a cell with asymmetric inhibition above the BF. (c) Response area of a cell with asymmetric inhibition below the BF.

We have examined the spatial distribution of these response area types in the AI of over 30 animals. Figure 3a illustrates one such map which was obtained using exclusively single-unit recordings in an experiment that lasted over 36 hrs. In all others, the maps were constructed from a mix of single unit and multiple units recordings [*Shamma, Fleshman and Wiser*, 1991]. Regardless of the recording method, the basic outline of the response area organization that emerges is as follows: at the center of AI, units respond with a narrow excitatory tuning curve, flanked by symmetric inhibitory side-bands. The response areas become more asymmetric away from the center. In one direction (caudally in the ferret AI), the inhibitory side-bands above the BF become relatively stronger. The opposite occurs in the other direction. These response types tend to organize along one or more bands that parallel the tonotopic axis (i.e., orthogonal to the isofrequency planes).

Organization of Selectivity to Direction of FM Tones

Cortical units were also tested using dynamic stimuli such as FM tones swept at different rates and intensities. The basic finding is that the selectivity of a cell's response to the direction of an FM tone correlates strongly with the symmetry of its response area. Specifically, cells with strong inhibition from frequencies above (below) the BF prefer upward (downward) moving sweeps. Thus, selectivity to FM direction is also mapped along the isofrequency planes of the AI, as demonstrated in Fig. 3b [*Shamma, Fleshman and Wiser*, 1991].

Organization of Responses to Spectrally Shaped Noise

Finally, several maps were also constructed from responses to spectrally shaped noise [*Shamma, Fleshman and Wiser*, 1991]. The stimuli consisted of spectral peaks of different symmetries as

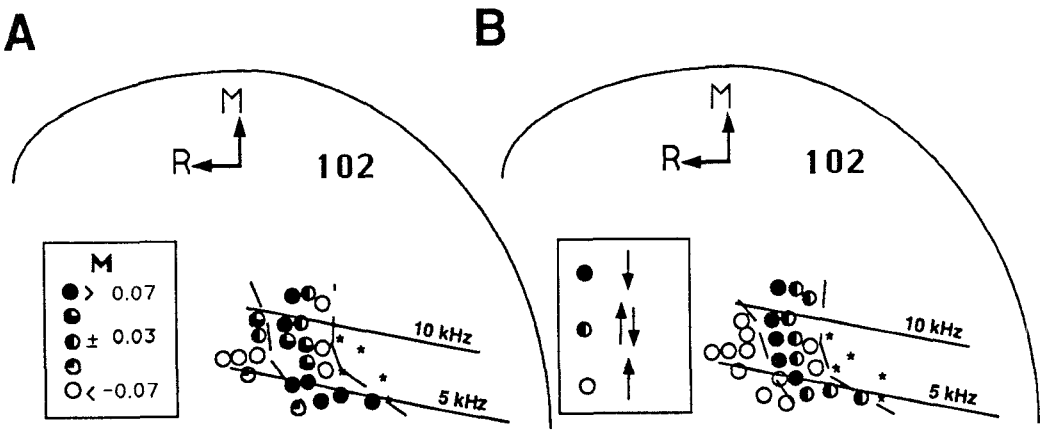


Fig. 3. (a) An example of the distribution of response area types, measured using the two-tone stimulus. The circles indicate the locations of the electrode penetrations along the isofrequency contours (shown schematically as solid straight lines); asterisks mark penetrations with weak auditory responses. Black (Clear) circles signify penetrations in which the response areas exhibit strong inhibition from below (above) the BF. Shaded penetrations are more symmetric. The response area symmetry index M is the percentage difference of responses around the BF. For details, please refer to [*Shamma, Fleshman and Wiser*, 1991].. the dashed lines delineate the approximate borders of the band within which the M measure changes *once* from extreme negative (clear circles) to extreme positive values (black circles). A key for the shading scheme used is shown on the left of the figure. The (M)edial and (R)ostral directions are indicated by the arrows; the arrow lengths represent 1/2 mm distances on the surface of the cortex. (b) Map of the topographic distribution of FM responses in AI. Map features and symbols are as in Fig. 3a. Black (clear) circles mark penetrations with selective responses to downward (upward) sweeps. Shaded penetrations are less selective.

illustrated in Fig. 4. Unit responses are consistent with the symmetry of their response areas. For instance, cells with strong inhibition from above the BF are most responsive to stimuli that contain least spectral energy above the BF, i.e., stimuli with the opposite asymmetry. Since response area symmetry is ordered along the AI, so is the local symmetry about the BF corresponding to the spectrum of the most effective stimulus (Fig. 4).

Discussion

What is the functional significance of these maps? One possible interpretation of the response area map is that it simply serves as the substrate needed to account for the organization of FM selectivity. Another possibility, however, is that other stimulus features besides FM selectivity are extracted and mapped simultaneously. This is the case for instance in the primary visual cortex where selectivity to the direction of edge motion (analogous to FM) and orientation are functionally linked [*Zeki and Shipp*, 1988]. The potential auditory equivalent of the visual edge orientation map is illustrated in the schematics of Fig. 4. Here, best responses to spectral peaks or edges of different symmetries are mapped systematically across the AI. The underlying functional principle of such a map is that, *the differential distribution of responses along the isofrequency planes of AI encodes a local measure of the shape of the acoustic spectrum - specifically, the locally averaged gradient of the spectrum* [*Shamma, Fleshman and Wiser*, 1991]. The response area maps in AI then can be described primarily as *gradient maps*. As such,

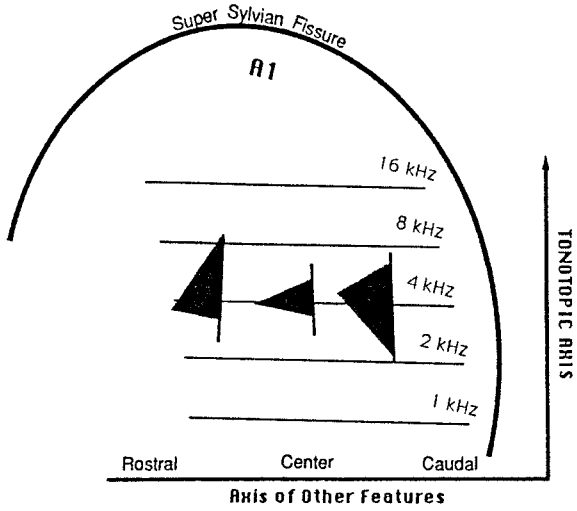


Fig. 4. Schematic of the AI responses to spectrally shaped stimuli. Cortical features are as in Fig. 1. Near the center, cells respond best to narrow symmetric spectral peaks, centered around the best frequency of the cell. Caudally, cells respond best to stimuli that extend to lower than BF frequencies, and lack energy above the BF. The opposite is true for rostrally located cells.

they can be viewed as the one dimensional analogue of the orientation columns of the visual cortex, since the orientation of a two-dimensional edge simply entails specifying its gradients in two directions.

PERCEPTION OF SPECTRAL PEAK SHAPES

The physiological experiments above suggest the existence of an explicit mapping for the shape of a spectral peak (Fig. 4). If so, this would possibly imply enhanced sensitivity and systematic trends in the perception of spectral gradients, and especially of spectral peak shapes. We addressed these issues directly in a series of experiments in which measurements were made of the sensitivity of human subjects to changes in the symmetry and bandwidth of spectral peaks.

Methods

The methods employed here follow closely those used in *Profile Analysis* experiments [*Green*, 1988]. The listener's task was to detect a change in the spectral shape of a *standard* complex waveform. In the frequency domain, the multicomponent *standard* consists of two portions, the base and the peak (Fig. 5, top). The base components are all equal in amplitude and they are added in phase to the peak components to form the peak shape (in dB scale).
The spectral shape change is achieved by changing the symmetry or bandwidth of the peak. These two factors define the shape of a peak uniquely as follows: (1) the bandwidth factor, BWF = 1/L – 1/R, where L and R are left and right slopes of the peak, respectively; (2) the symmetry factor, SF = (L + R) / (L – R). These factors are illustrated schematically in Fig. 6.

The discrimination task is to distinguish between the *standard*, which is unchanged over a block of trials, and the *signal*, which resembles the *standard* except for an adaptive change in peak shape in each trial. In all the experimental conditions, the *standard* and *signal* consist of 41 components equally spaced on a logarithmic scale between 200 Hz and 5 kHz. The peak is located at 1 kHz. Its amplitude is kept constant at a level (7.5 dB above the base) which allows it to be heard clearly.

Detection of Symmetry Changes

In these experiments, the bandwidth (BWF) is kept constant and a listener is forced to recognize

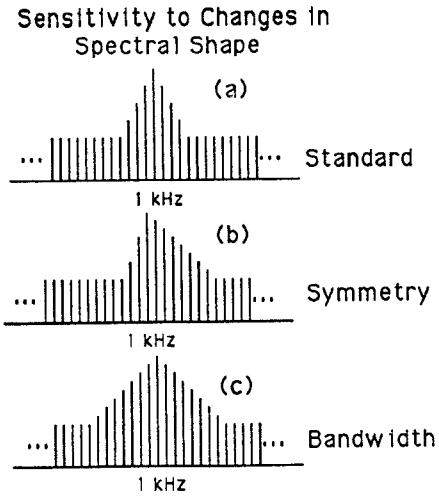


Fig. 5. (a) Complex waveform consists of an equal amplitude base and a peak added to it. Peak takes different symmetries (b) and bandwidths (c).

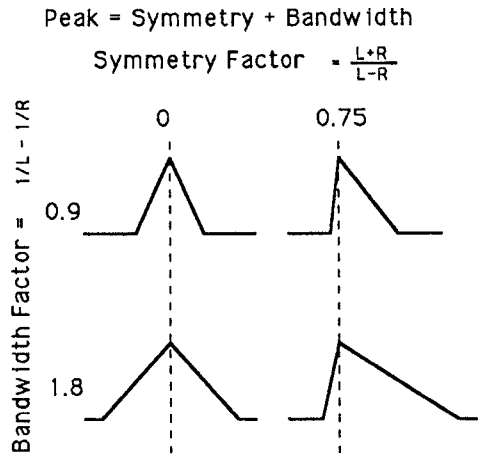


Fig. 6. The shape of a peak is determined by its symmetry and bandwidth factors. In each column, the peaks have the same symmetry but different bandwidths; in each row, the peaks have the same bandwidths, but different symmetry.

the *signal* based on a change in symmetry factor. The detection threshold is measured for *standard* peaks of different bandwidths (BWF = 0.1, 0.2, 0.13 and 0.4), and different symmetry factors (SF = 0, 0.1, 0.15, 0.2 and 0.4). Thus, a total of 20 tasks were run.

The average results for four subjects are presented in Fig. 7. The fundamental result that emerges is that the detection of a change in peak symmetry is independent of the peak shape of the standard. This is because: (1) at each BWF, the results for all five SF's match very closely, hence only the average value is plotted; (2) for all BWF's, the threshold remains relatively constant, i.e., the detection of symmetry is independent of bandwidth. Note that the detection threshold here ($\approx$ –9 dB using the measure defined in the figure legend) is comparable to that obtained previously in the detection of an edge [*Green*, 1988].

Detection of Bandwidth Changes

In these tests, we measure the listener's ability to detect a change in peak bandwidth, keeping SF constant. *Standards* at three different bandwidths (BWF = 0.1, 0.2, and 0.4) and five symmetries (SF = 0, 0.1, 0.15, 0.2, and 0.4) are used, i.e., a total of 15 different conditions were tested.

The averaged data over three subjects are given in Fig. 8a. First note that the detection

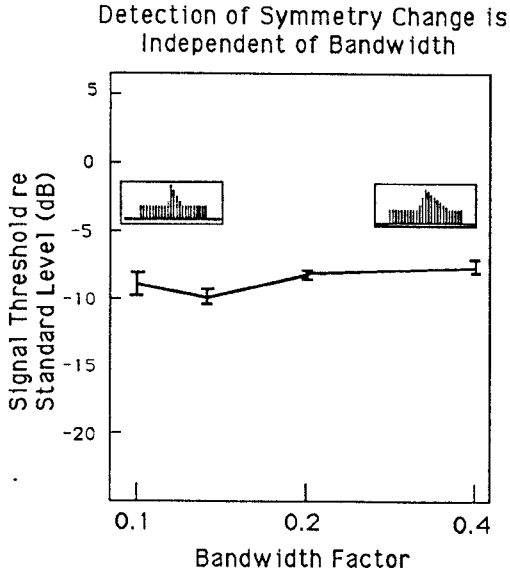


Fig. 7. Threshold for detection of symmetry change. The value along the ordinate is defined as: $20 \log \sqrt{\sum_{i=1}^{n} (\Delta a_i / a_i)^2}$, where Δa_i is the change in the amplitude of the i^{th} component at threshold, a_i is the amplitude of the i^{th} component in the *standard*, and $n = 41$. The bars represent the standard error. Note that thresholds do not depend on SF of *standard*, and hence all results from different SF's are collapsed into the curve shown.

thresholds are generally higher than before, and are relatively independent of SF. They are, however, *monotonically* dependent on the *standard's* bandwidth. The functional form of this dependence is seen more clearly in Fig. 8b, where the thresholds are averaged over all SF's and then plotted against BWF.

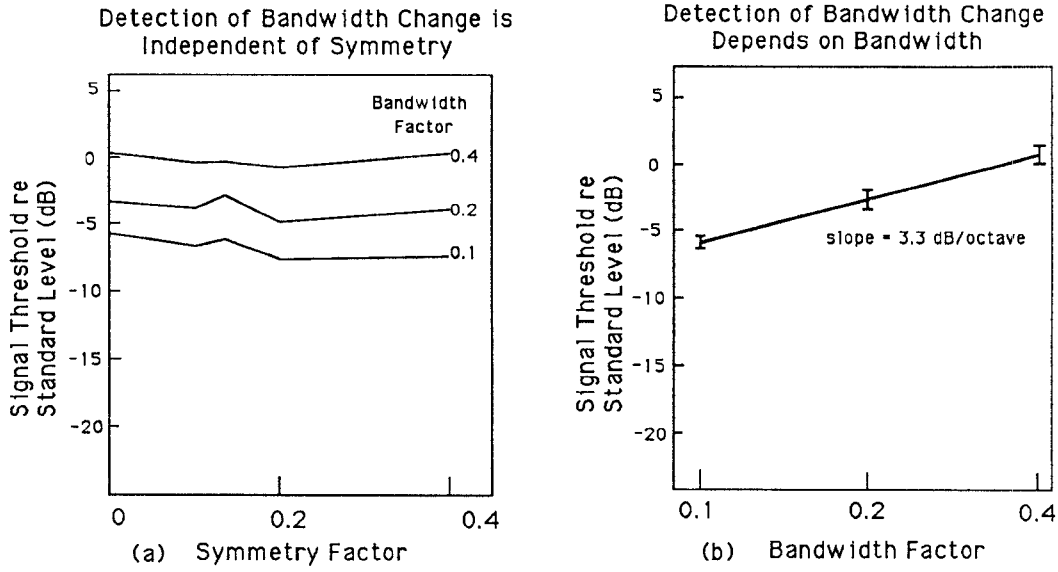


Fig. 8. (a) Detection threshold of bandwidth change as a function of SF for three different BWF's. (b) Average detection thresholds as a function of BWF of the *standard*.

Discussion

The basic finding that emerges from the psychoacoustical experiments described here is that the detection of a change in peak symmetry is both sensitive and independent of peak shape. Since peak symmetry is directly related to the average local gradient of the spectrum around the peak, these results can be viewed as the perceptual correlate of the the physiological gradient map. Furthermore, the independence of symmetry and bandwidth suggests that the gradient might be mapped physiologically more than once, each at a different scale of local averaging [*Shamma and Chettiar*, 1990]. While a physiological substrate for such a multiscale representation is yet unavailable in the auditory system, this idea is prevalent in the visual literature, and has a large body of psychophysical support behind it [*Mallat*, 1989].

ACKNOWLEDGMENTS

This work was supported by grants from the Air Force Office of Scientific Research, The Office of Naval Research and the National Science Foundation.

REFERENCES

Bernstein, L. R. and D. Green, Detection of simple and complex changes of spectral shape , *J. Acoust. Soc. Am.*, *82*, 1587–1592, 1987.

Blackburn, C. C. and M. B. Sachs, The representation of steady-state vowel sound /e/ in the discharge rate patterns of cat anteroventral cochlear nucleus, *J. Neurophysiology*, *63*, 1191–1212, 1990.

Green, D., *Profile Analysis*, Oxford Press, New York, 1988.

Hubel, D. and T. Wiesel, Receptive Fields, binocular interaction and functional architecture in the cat's visual cortex, *J. Physiol. (London)*, *160*, 106–154, 1962.

Imig, T. and H. Adrian, Binaural columns in the primary field (AI) of cat auditory cortex, *Brain Res.*, *138*, 241–257, 1977.

Kim, D. O., K. Parham, J. G. Sirianni and S. O. Chang, Spatial response profiles of posteroventral cochlear nucleus neurons and auditory nerve fibers in unanesthetized decerebrate cats: Response to pure tones, paper presented at in press, J. Acoust. Soc. Am., 1991.

Mallat, S., Multifrequency Channel Decompositions of Images and Wavelet Models, *IEEE Trans. Acoust. Speech Sig. Proc.*, *37(12)*, 2091–2110, 1989.

Reale, R. and T. Imig, Tonotopic organization of auditory cortex in the cat, *J. Comp. Neurol.*, *192*, 265–291, 1980.

Sachs, M. B. and E. D. Young, Encoding of steady state vowels in the auditory-nerve: representation in terms of discharge rate, *J. Acoust. Soc. Am.*, *66*, 470–479, 1979.

Shamma, S., J. Fleshman and P. Wiser, Response Area Organization in Primary Auditory Cortex, paper presented at in press J. Neurophysiol., 1991.

Shamma, S. A. and G. Chettiar, *A Functional Model of Primary Auditory Cortex: Spectral Orientation Columns*, System Research Center Tech. Rep. (TR 90-47), 1990.

Suga, N., The extent to which biosonar information is represented in the bat auditory cortex, in *Dynamic Aspects of Neocortical function*, edited by G. M. Edleman, W. E. Gall and W. M. Cowan, pp. 315–37, New York: Wiley, 1984.

Wenstrup, J., L. Ross and G. Pollak, Binaural response organization within a frequency band representation of the inferior colliculus: implications for sound localization, *J. Neuroscience*, *6*, 962–973, 1986.

Zeki, S. and S. Shipp, The functional logic of cortical connections, *Nature*, *335*, 311–317, 1988.