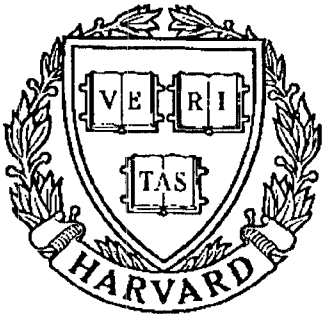
TECHNICAL RESEARCH REPORT

SYSTEMS RESEARCH CENTER



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

A Functional Model of Primary Auditory Cortex: Spectral Orientation Columns

by S. Shamma and G. Chettiar

TR 90-47

A Functional Model of Primary Auditory Cortex: Spectral Orientation Columns

Shihab A. Shamma[a] *and Geeth M. Chettiar*[b]

(a) Electrical Engineering Department, University of Maryland Institute for Advanced Computer Studies, and Systems Research Center. University of Maryland, College Park, MD 20742.

The Mathematical Research Branch, NIDDDK, National Institutes of Health, Bethesda, MD 20892.

(b) Electrical Engineering Department, University of Maryland, College Park, MD 20742.

Mailing Address: Shihab Shamma, Electrical Engineering Department, University of Maryland, College Park. MD. 20742.

Telephone numbers: (301) 454-6867 (Office). (301) 454-6942 (Lab).

ABSTRACT

A functional model of the primary auditory cortex is proposed based on physiological maps of the receptive field organization in ferret AI (Shamma, Fleshman and Wiser [1990]). Systematic changes in the excitatory and inhibitory portions of the receptive fields along the isofrequency planes are approximated by a difference of gaussians function with spatially changing parameters. We consider here only the response properties to *stationary* stimuli, i.e. those with non-varying spectra. The fundamental functional principle that emerges from the analysis of the model is that the primary auditory cortex encodes the shape of the acoustic spectrum in the distribution of its responses along the isofrequency planes. Specifically, it maps to each isofrequency plane a normalized measure of the locally averaged gradient of the input spectrum at that frequency. Physiological and psychoacoustical correlates and implications of these findings are explained. Parallels to the functional organization of the visual cortex are also discussed.

I. Introduction

The primary auditory cortex (AI) is essential for the perception and localization of sound. Its precise role in carrying out these functions, however, remains a mystery despite extensive knowledge gained from ablation experiments and from single and multi-unit recordings with various complex stimuli (Keidel and Neff [1975]; Neff, Diamond and Casseday [1975]; Schreiner and Urbas [1988]; Whitfield [1980]; Whitfield and Evans [1965]; Winter and Funkenstein [1973]). At present, the only two organizational features of AI that are firmly established are the spatially ordered tonotopic axis (Merzenich, Knight and Roth [1975]; Reale and Imig [1980]), and the alternating bands of binaural response properties that run perpendicularly to the isofrequency planes (Imig and Adrian [1977]; Middlebrooks, Dykes and Merzenich [1980]) (Fig.1). They are roughly analogous to the retinotopic maps and the ocular dominance columns of the primary visual cortex (Hubel and Wiesel [1962]).

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]). Ordered responses to more complex stimulus features, analogous to the orientation columns and direction of motion selectivity in the visual cortex, have been more difficult to find in AI. At present, only a few reports hint at the existence of such maps in AI (Mendelson et al. [1988]; Schreiner et al. [1988]).

This issue was addressed directly in a recent series of experiments in the ferret AI (Shamma, Fleshman and Wiser [1990]). The study explored the detailed organization of the receptive fields[1] of cells along the isofrequency planes. The aim was to establish whether any systematic changes in the balance of inhibitory and excitatory responses occur in cells along these contours and, if so, to determine the implications of these changes to responses to frequency-modulated (FM) tones and spectrally shaped noise stimuli. These response features are more complex

[1]The term *receptive field* of a cell is used here to denote its response as a function of frequency *at a particular intensity*. In a linear system, it would be called the *transfer function*. Since cortical cell responses do change with sound level, receptive fields are usually measured at several intensities.

than the determination of a single best frequency BF (tonotopicity) or the (binary) nature of a binaural interaction (e.g., an Excitatory-Excitatory or Excitatory-Inhibitory response). The receptive field of a cell represents, to first order, its transfer function, i.e., the way it *filters or processes* the input spectrum. Similarly, FM tones reveal information about the dynamic interplay between of the inhibitory and excitatory responses of the cell.

The basic findings of the above experiments can be summarized as follows (Fig.2):

- There seems to exist a spatially ordered change in the symmetry of the receptive fields in any given isofrequency plane in AI (Fig.2A). At the center, units respond with a narrow excitatory tuning curve at BF, flanked by narrow inhibitory side-bands. The receptive fields 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).

- Cell responses to spectrally shaped noise are consistent with the symmetry of their receptive fields. 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 receptive field symmetry is ordered along the AI. then so is the local symmetry about the BF of the spectral envelope of the most effective stimulus (Shamma, Fleshman and Wiser [1990]) (Fig.2B).

- The selectivity of a cell's response to the direction of an FM tone correlates strongly with the symmetry of its receptive fields. Specifically, cells with strong inhibition from frequencies above (below) the BF prefer upward (downward) moving sweeps (Fig.2C). Thus, selectivity to FM direction is also mapped along the isofrequency planes of the AI.

In this report, we discuss the significance of these findings to cortical function, and present a simplified mathematical model that seeks to answer the following questions: *What are the spatial distributions of cortical responses that result from such ordered receptive fields ? And what are the basic attributes of the stimulus that are mapped in such representations?* One

such attribute that is immediately apparent from the experimental data is the direction of an FM sweep. It is conceivable that the observed receptive field organization solely functions to generate the FM maps. However, it is also possible that other stimulus features are extracted and mapped simultaneously, as 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 second attribute of the acoustic spectrum that AI responses likely encode by their differential distribution along the isofrequency planes, is a local measure of the shape of the acoustic spectrum - specifically, the locally averaged *gradient* of the spectrum (Shamma, Fleshman and Wiser [1990]). This conjecture follows from the schematics of Fig.2B where best responses to spectral peaks or edges of different symmetries are mapped systematically across the AI. The significance of such a map stems from its enhancement and explicit representation of such perceptually important features as the shape of spectral peaks, edges, and the spectral envelope (Assmann and Summerfield [1988]; Klatt [1982]). Furthermore, this gradient map can be viewed as a 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.

In what follows, we shall elaborate on these arguments using a simplified model of the receptive field organization of AI. First, we review the most salient features of the receptive fields across AI as seen in the experimental data. Next, the receptive fields are approximated by gaussian functions and an analysis of the resulting cortical response patterns is carried out. Finally, we examine the implications of the cortical models and its outputs to synthetic and natural stimuli.

II. The Salient Features of Receptive Field Organization in AI

The receptive fields in AI were measured in a series of mapping experiments described in detail in (Shamma, Fleshman and Wiser [1990]). Briefly, cells sampled along the isofrequency contours were driven by a two-tone stimulus in which the first tone, $T1$, spanned a wide range of frequencies about the BF (± 1 or ± 2 octaves). The second tone, $T2$, was fixed at BF to

provide a level of response against which the inhibition can be detected. $T1$ usually preceded $T2$ by a small interval (30 ms) so that the responses to each tone can be visually segregated in the rasters. An example of the responses of a cell with narrow and symmetric lateral inhibition is shown in the middle raster of Fig.3. The symmetry of the inhibition can also be seen in the combined spike count curve computed in the window 100–180 ms and illustrated below the raster. This response curve as a function of frequency is what is called here the *receptive field* of the cell. For each cell, two or more such curves were measured at various intensities. In most cells, the gross features of the receptive field (e.g., the strength and symmetry of the inhibition) do not depend critically on the relative levels of the tones or the inter-tone delay (Shamma, Fleshman and Wiser [1990]).

Systematic changes in the symmetry of the receptive fields emerge along the isofrequency contours of AI (Shamma, Fleshman and Wiser [1990]). Three examples covering the range of response symmetries observed are shown in the rasters of Fig.3. For each raster, an *index (M)* is computed in order to describe concisely the overall character of the receptive field. The index reflects the net balance of inhibition and excitation around the BF of the cell and is computed from the receptive field curve as follows:.

$$M = \frac{R_{>BF} - R_{<BF}}{R_{>BF} + R_{<BF}},$$

where $R_{>BF}$ and $R_{<BF}$ are the total number of spikes for all frequencies above and below the BF, respectively.

Based on this index, it is possible to classify the continuum of cell response types into one of three catagories: Cells with strong inhibition from above the BF (large negative M values, e.g. < -0.07), cells with nearly symmetric inhibition (M values near zero), and cells with strong inhibition from low frequencies (large positive M values, > 0.07) (Shamma, Fleshman and Wiser [1990]). For each of these types, we have obtained an *average or typical* raster and receptive field by collapsing the responses from all cells that belong to the group, across many animals. Fig. 4 illustrates the resulting rasters and curves for the three groups, each with two frequency ranges about the BF. Note that the frequency axes are now labeled relative to a nominal BF since the averaged responses have different BFs. Note also that the amplitude

of the excitatory response at BF is artificially depressed relative to the "background" rate far from the BF (i.e., at either end of the curve). This is an inevitable consequence of the use of $T2$ to raise artificially the response baseline so as to detect the inhibition.

The rasters and receptive fields of Fig. 4 display two of the three basic features of the way the receptive fields change along an isofrequency plane in AI (Shamma, Fleshman and Wiser [1990]). These are:

1. The changes in the symmetry of the receptive fields can be viewed conceptually as the consequence of a gradual shift or misalignment of a broadly tuned inhibition (total width 2 - 2.5 octaves) overlapping a narrower excitatory response area around the BF (1/2 - 3/4 octave wide). At the center of AI, both the excitatory and inhibitory inputs are aligned, and only relatively narrow symmetric inhibitory surrounds are uncovered ($\approx$ 3/4 octave on either side). Moving towards the edges of AI, the inhibition shifts lateraly causing an asymmetric spread and strengthening of the inhibition on one side of the BF, and a general weakening on the other.

2. The movement of the inhibitory surround is associated with a slight opposite shift of the BF and a moderate broadening in the bandwidth of the excitatory tuning curve (by less than a 1/4 octave). These changes can be seen clearly in the superimposed plots in Fig. 5., where the center (symmetric) receptive field is compared directly to each of the two asymmetric curves.

3. The third consistent finding in our experiments - one that is difficult to substantiate in single unit records like Fig. 4 - is the gradual weakening of the excitatory evoked potentials towards the borders of AI. This trend may reflect a decreasing density of responsive cells or a simple threshold elevation. In our data, no obvious systematic increase in the threshold of responses of isolated units could be discerned. In the cat AI, however, the same trend has been linked to such threshold elevation (Schreiner et al. [1988]). In either case, the functional implication is that the amplitude of the excitatory portion of the receptive fields becomes smaller for asymmetric cells.

Taking these three basic features into account, it is possible to formulate a simple recipe to generate the receptive fields of the primary auditory cortex.

III. Modelling the Receptive Fields Along the Isofrequency Planes of AI

In order to determine theoretically the functional significance of the receptive field organization in AI, it is necessary to develop a mathematically tractable description of these fields, and of the way they vary as a function of location. One choice that is particularly convenient is based on a combination of two gaussian functions, $G(\mu_e, \sigma_e; f)$ and $G(\mu_i, \sigma_i; f)$, representing the excitatory and the inhibitory inputs, respectively. The mean and standard deviation parameters are represented by μ and σ, respectively; f, frequency, is the independent variable. The two functions have the following explicit forms:

$$G(\mu_e, \sigma_e; f) = \frac{1}{\sqrt{2\pi\sigma_e^2}} \exp -\frac{(f-\mu_e)^2}{2\sigma_e^2} \equiv G_e$$

and,

$$G(\mu_i, \sigma_i; f) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp -\frac{(f-\mu_i)^2}{2\sigma_i^2} \equiv G_i.$$

Using these functions, we can construct a series of model receptive fields at different locations (x) along a specific isofrequency contour (f_o) as follows (Fig. 6):

- The excitatory gaussian curve is centered at f_o by setting $\mu_e = f_o$. This parameter remains fixed for all receptive fields along the contour.

- The value of σ_e is chosen to reflect the approximate width of the excitatory portion of the receptive field. Since this width varies as a function of location (x) along the isofrequency plane (2^{nd} feature above and Fig.5), we take σ_e to be a function of x. $\sigma_e(x)$ is then smallest near the center of AI ($x = 0$) where the tuning is sharpest, and increases towards the edges ($x \equiv x_m, -x_m$; Upper row in Fig.6). Note that the increasing bandwidth towards the edges is associated with a decreasing amplitude, thus satisfying the 3^{rd} feature of the physiological data.

- The bandwidth of the inhibition is assumed constant and larger than the bandwidth of the excitation along the isofrequency plane (i.e., $\sigma_i > \sigma_e(x)$ for all x). Since the inhibition moves relative to the excitation, then μ_i is a function of x. At the center of AI ($x = 0$), the inhibition is aligned with the excitation, i.e., $\mu_i(0) = \mu_e(= f_o)$. Towards the edges of AI, $\mu_i(x)$ increasingly deviates from μ_e (middle row in Fig.6).

- Finally, the inhibitory gaussian curve is subtracted from the excitatory curve to obtain the receptive field (RF) at each location x along the isofrequency plane f_o (Last row in Fig.6):

$$RF(x, f) = G(\mu_e, \sigma_e(x); f) - G(\mu_i(x), \sigma_i; f)$$

where $\mu_i(x) = \mu_i(0) + \delta\mu_i(x)$ and $\sigma_e(x) = \sigma_e(0) + \delta\sigma_e(x)$.

There are four parameters that need to be determined from the physiological data: $\sigma_e(0)$ and σ_i, representing the widths of the excitation and inhibition at the center of AI, respectively; $\delta\sigma_e(x)$, representing the functional form of the broadening of the excitatory bandwidth; and $\delta\mu_i(x)$, describing the relative shift of the inhibition. The first two parameters can be roughly estimated from the symmetric receptive field in Fig. 4 ($\sigma_e(0) \approx 4$, and $\sigma_i \approx 8$). However, we have no data that can provide a reliable indication of the exact way the parameters change with distance across the AI since the physiological maps vary considerably across animals. We have therefore assumed the following simple odd and even functions for $\delta\mu_i(x)$ and $\delta\sigma_e(x)$, respectively:

For the inhibitory shifts: $\delta\mu_i(x) = (a/x_m) \cdot x$.

For the excitatory broadening: $\delta\sigma_e(x) = (b/x_m^2) \cdot x^2$.

The rate constants (a/x_m and b/x_m^2) are chosen to generate approximately the asymmetric receptive fields of Fig. 4 near the caudal and rostral edges of AI. Fig.7 illustrates the match between the physiological curves and the model receptive fields using the following parameter values:

$\sigma_e(0) = 4.0$, $\sigma_i = 8.0$, $\delta\mu_i(x) = 0.15x$, $\delta\sigma_e(x) = 0.008x^2$

The profiles compare well except for the artificial elevation of the baseline in the physiological data due to the use of the second tone $T2$ discussed earlier.

IV. Analysis of Receptive Field Organization in AI

In this section, we carry out an approximate analysis of the mathematical model of AI. The analysis aims to confirm the intuitive interpretation of the function of AI outlined in the introduction - that it encodes the locally averaged gradient of the acoustic spectrum along its isofrequency planes. We shall also seek to determine the significance of the different receptive field parameters in the formation of the response maps. The analysis is approximate in that only the first- and second-order terms of the taylor expansions of the receptive field models are considered. As such, the expressions derived will be inaccurate for certain ranges of the parameters and for some input spectra, as we shall elaborate.

The receptive field ($RF(x,f)$) at location x along the isofrequency f_o is described in the model by the difference of gaussians:

$$RF(x,f) = G(\mu_e,\sigma_e;f) - G(\mu_i,\sigma_i;f) \equiv G_e - G_i. \tag{1}$$

Since both the σ_e and μ_i vary as a function of x, we write:

$$\begin{aligned}\sigma_e &= \sigma_e(0) + \delta\sigma_e(x) \equiv \sigma_o + \delta\sigma_e(x)\\ \mu_i &= \mu_i(0) + \delta\mu_i(x) \equiv \mu_e + \delta\mu_i(x).\end{aligned}$$

We now expand the two gaussians about their respective centers up to their second-order terms:

$$G_e = G(\mu_e,\sigma_o+\delta\sigma_e;f) \approx G(\mu_e,\sigma_o;f) + \delta\sigma_e \left.\frac{\partial G_e}{\partial\sigma}\right|_{\sigma_e=\sigma_o} + \frac{(\delta\sigma_e)^2}{2}\left.\frac{\partial^2 G_e}{\partial\sigma^2}\right|_{\sigma_e=\sigma_o}$$

$$G_i = G(\mu_e+\delta\mu_i,\sigma_i;f) \approx G(\mu_e,\sigma_i;f) + \delta\mu_i \left.\frac{\partial G_i}{\partial\mu}\right|_{\mu_i=\mu_e} + \frac{(\delta\mu_i)^2}{2}\left.\frac{\partial^2 G_i}{\partial\mu^2}\right|_{\mu_i=\mu_e}.$$

In order to simplify the notation, we shall drop in the remainder of the text the subscripts evaluating the derivatives, while assuming that all derivatives are evaluated as indicated in the equations above. Using the following two properties of the derivatives of the gaussian curve:

$$\frac{\partial G}{\partial\sigma} = \sigma\frac{\partial^2 G}{\partial\mu^2},$$

$$\frac{\partial^2 G}{\partial \sigma^2} = \frac{\partial^2 G}{\partial \mu^2} + \sigma^2 \frac{\partial^4 G}{\partial \mu^4},$$

we obtain:

$$\begin{aligned} G_e &\approx G(\mu_e, \sigma_o; f) + \delta\sigma_e\sigma_o \frac{\partial^2 G_e}{\partial \mu^2} + \frac{(\delta\sigma_e)^2}{2}\frac{\partial^2 G_e}{\partial \mu^2} + \frac{(\delta\sigma_e\sigma_o)^2}{2}\frac{\partial^4 G_e}{\partial \mu^4} \\ G_i &\approx G(\mu_e, \sigma_i; f) + \delta\mu_i \frac{\partial G_i}{\partial \mu} + \frac{(\delta\mu_i)^2}{2}\frac{\partial^2 G_i}{\partial \mu^2}. \end{aligned}$$

The receptive field in Eq.1 becomes approximately:

$$\begin{aligned} \widetilde{RF}(x,f) &= (G(\mu_e, \sigma_o; f) - G(\mu_e, \sigma_i; f)) + (\delta\sigma_e\sigma_o + \frac{(\delta\sigma_e)^2}{2})\frac{\partial^2 G_e}{\partial \mu^2} + \frac{(\delta\sigma_e\sigma_o)^2}{2}\frac{\partial^4 G_e}{\partial \mu^4} \\ &\quad -\delta\mu_i \frac{\partial G_i}{\partial \mu} - \frac{(\delta\mu_i)^2}{2}\frac{\partial^2 G_i}{\partial \mu^2} \end{aligned}$$

Since $\frac{\partial G}{\partial \mu} = -\frac{\partial G}{\partial f}$ and $\frac{\partial^2 G}{\partial \mu^2} = \frac{\partial^2 G}{\partial f^2}$, the derivatives in the above equation can be rewritten as:

$$\begin{aligned} \widetilde{RF}(x,f) &= G(\mu_e, \sigma_o; f) - G(\mu_e, \sigma_i; f)) + \underbrace{\delta\mu_i}_{o(x)} \frac{\partial G_i}{\partial f} + \underbrace{(\delta\sigma_e\sigma_o + \frac{(\delta\sigma_e)^2}{2})}_{-e_1(x)} \frac{\partial^2 G_e}{\partial f^2} \\ &\quad \underbrace{-\frac{(\delta\mu_i)^2}{2}}_{-e_2(x)} \frac{\partial^2 G_i}{\partial f^2} + \underbrace{\frac{(\delta\sigma_e\sigma_o)^2}{2}}_{e_3(x)} \frac{\partial^4 G_e}{\partial f^4} \end{aligned}$$

Let:

$$\begin{aligned} o(x) &= \delta\mu_i(x) = 0.15x; \\ e_1(x) &= -\delta\sigma_e(x)\sigma_o - \frac{(\delta\sigma_e(x))^2}{2} \\ &= -0.008x^2(4) - \frac{(.008x^2)^2}{2} = -0.032x^2 - 0.000032x^4 \\ e_2(x) &= \frac{(\delta\mu_e(x))^2}{2} \\ &= \frac{(0.15x)^2}{2} = 0.01125x^2 \\ e_3(x) &= \frac{-(\delta\sigma_e(x)\sigma_o)^2}{2} \\ &= \frac{(.008x^2(4))^2}{2} = 0.000512x^4 \end{aligned}$$

So we have:

$$\widetilde{RF}(x,f) = (G(\mu_e, \sigma_o; f) - G(\mu_e, \sigma_i; f)) + o(x)\frac{\partial G_i}{\partial f} - e_1(x)\frac{\partial^2 G_e}{\partial f^2} - e_2(x)\frac{\partial^2 G_i}{\partial f^2} + e_3(x)\frac{\partial^4 G_e}{\partial f^4}$$

Distribution of Cortical Activity

In order to compute the approximate distribution of cortical activity $\tilde{A}(x, f_o)$ along the isofrequency plane f_o, for a given input spectrum $I(f)$, we convolve the input with the receptive field expression above. In doing so, we imply that the receptive fields are linear and shift-invarient - two assumptions that are likely to be only approximately true:

$$\begin{aligned}
\tilde{A}(x,f) &= \widetilde{RF}(x,f) * I(f) \\
&= [G(\mu_e,\sigma_o;f) - G(\mu_e,\sigma_i;f) + o(x)\frac{\partial G_i}{\partial f} - e_1(x)\frac{\partial^2 G_e}{\partial f^2} - e_2(x)\frac{\partial^2 G_i}{\partial f^2} e_3(x)\frac{\partial^4 G_e}{\partial f^4}] * I(f) \\
&= (G(\mu_e,\sigma_o;f) - G(\mu_e,\sigma_i;f)) * I(f) + o(x)\frac{\partial I}{\partial f} * G_i - e_1(x)\frac{\partial^2 I}{\partial f^2} * G_e - e_2(x)\frac{\partial^2 I}{\partial f^2} * G_i \\
&\quad + e_3(x)\frac{\partial^4 I}{\partial f^4} * G_e,
\end{aligned}$$

where use is made of the fact that $\frac{\partial G}{\partial f} * I(f) = \frac{\partial I}{\partial f} * G$. The last expression can be notationally simplified further to be:

$$\tilde{A}(x,f_o) = \underbrace{(G(\mu_e,\sigma_o;f) - G(\mu_e,\sigma_i;f)) * I}_{(1)} + \underbrace{o(x)\bar{I}_i^{(1)}}_{(2)} + \underbrace{e_1(x)\bar{I}_e^{(2)-} + e_2(x)\bar{I}_i^{(2)-}}_{(3)} + \underbrace{e_3(x)\bar{I}_e^{(4)}}_{(4)} \quad (2)$$

where $\bar{I}_i^{(1)} \equiv \frac{\partial I}{\partial f} * G_i$ is the locally averaged gradient of the input spectrum, $\bar{I}_e^{(2)-} \equiv -\frac{\partial^2 I}{\partial f^2} * G_e$ and $\bar{I}_i^{(2)-} \equiv -\frac{\partial^2 I}{\partial f^2} * G_i$ are the inverted locally averaged second-derivatives of the input spectrum (both are large positive near sharp spectral peaks and edges, and small or negative otherwise), and $\bar{I}_e^{(4)} \equiv \frac{\partial^4 I}{\partial f^4} * G_e$ is the locally averaged fourth-derivative of the input spectrum (it is large positive near sharp spectral peaks and edges).

V. Interpreting the Response Distribution

Equation 2 describes the response patterns across the surface of AI for a given spectral pattern $I(f)$. Most important for our purposes is the interpretation of the differential distribution along the isofrequency plane, i.e., the dependence of $\tilde{A}(x, f_o)$ on the variable x. There are four different types of terms in Eq.2:

1. The first term is independent of x and hence makes no contribution to the *differential* responses along the isofrequency contour. Instead, it provides a constant bias or background

level of activity which is modified by the other terms. It is nevertheless very important in that it effectively selects or weights more heavily the representation of regions of high curvature in the input spectrum $I(f)$. This is because the difference of gaussians operator $(G(\mu_e, \sigma_o; f) - G(\mu_e, \sigma_i; f))$ qualitatively measures the inverse of the second-derivative of the input spectrum at f_o. Consequently, spectral regions of sharp peaks and edges will produce large positive bias levels, whereas spectral valleys produce negative biases and hence are suppressed in the cortical outputs. Fig.8 illustrates the cortical outputs due to this term alone for a three peak input spectrum.

2. The second term in the equation is sensitive to the sign and magnitude of the locally averaged first-derivative of the spectrum at f_o. The form of this dependence is explained in Fig.9 for three types of spectral peaks. The plots on the left illustrate how $\bar{I}_i^{(1)}$ depends on the *local symmetry* of the peak, i.e., on the net balance of spectral slopes around f_0. Since the function $o(x)$ is assumed to be a straight line ($o(x) = 0.15x$), then $\bar{I}_i^{(1)}$ simply acts to modify the slope of this line and hence the distribution of cortical activity around $x = 0$. This is illustrated in the plots to the right of Fig.9.

3. The third term is an even function of x, and is weighted by the magnitude of the second-derivative of the input spectrum at f_o. Its contribution to the cortical output is explained in Fig.10. The plots to the left depict the values of $\bar{I}_e^{(2)-}$ or $\bar{I}_i^{(2)-}$ near a typical spectral peak at f_o. Since $\bar{I}_e^{(2)-} \geq \bar{I}_i^{(2)-}$ (because the excitatory gaussian curve is more compact), and since the curvature of $e_1(x)$ is larger than that of $e_2(x)$, then this term is a concave function of x with a single maximum at $x = 0$. This is illustrated in the plots to the right of Fig.10. Note that the location of this maximum may shift if we add the contribution of the second (odd) term of Eq.2 (Fig.9). Specifically, the maximum shifts to the right for positive $\bar{I}_i^{(1)}$, and vise versa. The amount of this shift (not its direction), however, is also inversely related to the degree of concavity of the even function.

4. The fourth term is relatively small. Its contribution to cortical activity occurs mostly near the edges ($|x| > 10$) for input spectral peaks that are rather sharp, i.e., they posses sizable averaged 4^{th} derivatives.

The close correspondence between the approximate expression of $\tilde{A}(x, f_o)$ (Eq.2) and the outputs obtained from computing with the original receptive fields is demonstrated in Fig.11 for the case of a synthetic pattern and a natural speech spectrum.

The Spectral Shape Factor

A fundamental feature the distribution of cortical activity along the isofrequency plane f_o is the location of its maximal response, x_{max}. Therefore, a theory of the functional organization of AI must answer the following question: *What specific features of the input spectrum determine the location of maximal response along an isofrequency plane?*

To answer this question, we first derive an estimate of x_{max} directly from the expression for $A(x, f_o)$:

$$\begin{aligned} A(x, f_o) &= RF(x, f) * I(f) \\ &= (G_e(\mu_e, \sigma_e; f) - G_i(\mu_i, \sigma_i; f)) * I(f) \end{aligned}$$

by setting its derivative with respect to x to zero:

$$\begin{aligned} \frac{\partial A(x, f_o)}{\partial x} &= (\frac{\partial G_e}{\partial \sigma_e}\frac{\partial \sigma_e}{\partial x} - \frac{\partial G_i}{\partial \mu_i}\frac{\partial \mu_i}{\partial x}) * I(f) \\ &= (\sigma_e \frac{\partial G_e^2}{\partial f^2}\frac{\partial \sigma_e}{\partial x} + \frac{\partial G_i}{\partial f}\frac{\partial \mu_i}{\partial x}) * I(f) \\ &= 0. \end{aligned}$$

Substituting $\delta\mu_i = a'x$ and $\delta\sigma_e(x) = b'x^2$ into the equation above:

$$\begin{aligned} (2b'^2 x^3 \frac{\partial G_e^2}{\partial f^2} + a'\frac{\partial G_i}{\partial f}) * I(f) &= 2b'^2 x^3 (\frac{\partial G_e^2}{\partial f^2} * I(f)) + a'(\frac{\partial G_i}{\partial f} * I(f)) \\ &= -2b'^2 x^3 \bar{I}_e^{(2)-} + a'\bar{I}_i^{(1)} \\ &= 0. \end{aligned}$$

Solving for x, we obtain:

$$x_{max} = (\frac{a'}{2b'^2})^{1/3} \cdot (S)^{1/3} \tag{3}$$

where S, the *spectral shape factor*, is defined as

$$S \equiv \frac{\bar{I}_i^{(1)}}{\bar{I}_e^{(2)-}} \tag{4}$$

$$\equiv \left(\frac{\partial(I * G_i)}{\partial f}\right) / \left(\frac{\partial^2(-I * G_e)}{\partial f^2}\right) \tag{5}$$

These expressions describe the location of the maximal response in relation to the spectral pattern presented to the cortex. Since $\tilde{A}(x, f)$ is large only near spectral regions of fast curvature, then x_{max} basically reflects the character of spectral peaks and edges. S can be thought of as a *normalized* measure of the *local* average gradient of the input spectrum. The meaning of the two qualifiers (*normalized* and *local*) is made apparent by the definitions in (4) and (5) above. First in (4), the gradient (and hence the cortical map) is normalized by $\bar{I}_e^{(2)-}$. In the case of spectral peaks, this is roughly inversely proportional to the bandwidth. An intuitive understanding of the significance of the ratio $\bar{I}_i^{(1)}/\bar{I}_e^{(2)-}$ can be gained from the simple example in Fig.12. Here a schematic spectral peak at f_o is shown in relation to the receptive field. The peak has a well localized curvature ($I^{(2)-}$) and is surrounded by relatively constant (L)eft and (R)ight slopes *that overlap with the inhibitory portion of the receptive field* (Fig.12A). In this case, the S factor evaluated at f_o approximately varies as:

$$\begin{aligned} S(f_o) &= \frac{\int I^{(1)}(f) \cdot G_i(f) df}{\int I^{(2)-}(f) \cdot G_e(f) df} \approx \frac{0.5 \cdot (L + R)}{G_e(f_o) \cdot (L - R)} \\ &\sim \frac{L + R}{L - R}. \end{aligned}$$

This last expression suggests that S captures a *relative* measure of the shape of the spectral pattern, one that is insensitive to its bandwidth. For instance, the S value for a skewed spectral peak remains unchanged if the pattern is slightly dilated (Fig.12B). Similarly, all edges are signified by the same S value regardless of the steepness of their sides (Fig.12B). Note that according to this ratio, (one-dimensional) shapes vary from symmetric peaks ($\frac{L-R}{L-R} = 0$), to peaks of various symmetries ($|\frac{L+R}{L-R}| < 1$), to edges ($|\frac{L+R}{L-R}| \approx 1$), to nearly straight lines ($|\frac{L+R}{L-R}| \gg 1$).

The above estimates of S are of course only valid when the width of the spectral peak and of the receptive field are of comparable *scales.* This brings us to the meaning of the second qualifier made earlier, that of the *locality* of the mapping. As is evident from expression (5) for the S factor, x_{max} does not reflect the features of the original input spectrum $I(f)$, but rather of the spectrum locally smoothed by the gaussians. Consequently, *the scale of the spectral features*

detected and mapped by the cortex is primarily determined by the scale of the receptive fields. For example, consider the schematic spectral pattern of the two symmetric peaks in Fig.12C. A cortical map with narrow receptive fields will view the peak at f_o as a locally symmetric peak. In contrast, receptive fields that are broad relative to the peak will detect the presence of neighboring peaks and hence view it as highly asymmetric. In this sense, different scales offer complementary perspectives of the input spectral pattern ranging from detailed spectral peaks and inflections, to coarse features such as the overall shape of the spectral envelope.

Physiological receptive fields of all three types of symmetry vary considerably in their bandwidths (Shamma, Fleshman and Wiser [1990]). In formulating this model, we have used the receptive fields in Fig.5 which were of intermediate widths being the averages of many instants of varying widths. Instead, it is possible to recompute the model outputs for various dilations of the receptive fields, i.e., effectively to create a third response dimension which reflects the scale of the spectral features mapped. This multiscale representation of pattern features is already prevalent in the visual literature, and has a large body of psychophysical support behind it (Levine [1985]; Mallat [1989]).

VI. Examples of Cortical responses to Synthetic and Natural Acoustic Spectra

In this section we illustrate briefly examples of maps produced by various spectral patterns. In Fig.13, a series of synthetic patterns demonstrate two basic features: The shifts in x_{max} reflecting the changing symmetry of spectral peaks, and the relevance of the relative scales of the input features and the receptive fields. In the *leftmost* plot (Fig.13A), the input pattern contains three peaks of different symmetries. The center peak is symmetric, and hence the x_{max} is centered. x_{max} shifts to the right for the peak that spreads towards the low-frequency; The opposite occurs for the other peak. In the series of plots to the right (Fig.13B-D), all input peaks are symmetric and of equal amplitude, but the spacing between them is progressively decreased. When far apart relative to the width of the receptive fields, the interaction is minimum and hence three separate centered regions of activity appear (Fig.13B). When the peaks are closer together, the cortical output begins to merge and reflect more the envelope of the pattern, rather than its detailed peaks (Fig.13C). Consequently, the maximal regions of activity occur

now near the edges of the pattern (not near the center peak) and, furthermore, they are shifted towards the left and right of center to reflect the local symmetry of the edges. Finally, when the peaks are very close, the pattern is treated as one broad symmetric peak (fig.13D).

The patterns in Fig.14 are spectra of naturally spoken speech phonemes. They are computed using a cochlear model followed by lateral inhibitory networks, as described in detail in (Shamma [1989]). All details of these cortical outputs can be interpreted in light of the two features described above in Fig.13.

VII. Discussion

The primary auditory cortex encodes the shape of the acoustic spectrum in the distribution of its responses along the isofrequency planes. Specifically, it maps to each isofrequency plane a normalized measure of the locally averaged gradient of the input spectrum at that frequency. This is the functional principle that emerges from the analysis of a model of the receptive field organization in AI.

Relation to Cortical Physiology

The physiological data from the two-tone experiments were critical in the construction of the receptive fields of the model. However, of more direct correspondence to the functional principle stated above are the physiological mappings obtained using spectrally shaped noise stimuli (Shamma, Fleshman and Wiser [1990]) (Fig.2B). In these experiments, ordered maps of the *spectral shape factor (S)* were obtained along the isofrequency planes using a symmetric peak and two edges of opposite symmetries. These maps are consistent with the receptive field (two-tone) maps in that they always overlap and, in a given animal, share the same trends and dimensions (Shamma, Fleshman and Wiser [1990]). This is of course completely predicted in the model since the spectral shape mapping is a direct consequence of the receptive field organization.

The primary auditory cortex is only the last of a long chain of processing stages in the mammalian auditory system. We have at present no knowledge of the exact form of its inputs from the medial geniculate body, nor of its functional anatomy. It is for instance uncertain

whether the receptive field organization is of cortical or pre-cortical origins. It is also unclear what features are encoded *prior* to the cortex along the isofrequency planes which seem to exist as early as the cochlear nuclei (Bourk, Mielcarz and Norris [1981]). In all the experiments and theoretical analysis presented here, we have always visualized the cortical input to be the unidimensional spectral pattern produced essentially by the cochlea (and possibly the cochlear nucleus (Shamma [1989]). In this case, pre-cortical isofrequency planes are assumed to serve unrelated functions, such as binaural processing (Wenstrup, Ross and Pollak [1986]) or pitch mappings (Schreiner and Langner [1988]). However, it is entirely plausible that cortical input patterns are considerably more complex, and that in using stationary single tones and spectral patterns only certain static dimensions of AI processing are revealed. Instead, by using dynamic spectral patterns such as FM tones, it is possible to establish other mappings that coexist with the receptive field organization (e.g., the FM maps in Fig.2C). Another theoretical possibility is that the cortical input is a two dimensional pattern of activity, e.g., similar to Lickleider's correlograms (Slaney and Lyon [1990]), and that cortical receptive fields should be defined not only with respect to frequency, but also in terms of the feature appropriate to the second dimension. However, until more relevant physiological data are available, the tonotopic axis remains the only definite pre-cortical axis of monaural processing.

Relation to Psychoacoustics

Spectral shape recognition and discrimination has been addressed in psychoacoustical studies for many decades, albeit indirectly, as for instance in studies of musical timbre and speech perception. The few direct studies that exist are the *profile analysis* methods that Green *et al.* pioneered (Green [1988]), and the phonetic distance measurements (Assmann and Summerfield [1988]; Klatt [1982]). Profile analysis experiments so far have not dealt with spectral shape discrimination in the sense used in this paper. Rather, they have been mostly limited to single tone detection in complexes - a stimulus which probably corresponds to narrow symmetric peaks. The methods of profile analysis, however, can be directly extended to probe the perceptual relevance of peak asymmetries and other spectral shape features.

A second type of experiments have sought to determine the spectral metric, and hence

the spectral features, that listeners use in classifying or ordering different speech phonemes. Since only complex vowel spectra were used, it is difficult to interpret the results in terms of the more elementary spectral shape features discussed in this paper. Nevertheless, a consistent finding of these experiments is that in matching different spectra, listeners always weight disproportionately the regions of high curvature (e.g., the formants). Furthermore, the best distance metric that explains the resulting matches is one that uses a combination of the first and second derivatives of the spectrum (Assmann and Summerfield [1988]). Both findings are in agreement with the conclusions of the physiological data and of the cortical model. It is quite feasible to address more directly our findings with this experimental technique by simply limiting the complexity of the stimulus spectra to single peaks and edges of various symmetries.

Relation to the Visual System

The visual system presents direct analogs to the physiological data and theoretical conclusions discussed here. For instance, receptive fields of various asymmetries abound in the primary visual cortex (Jones and Palmer [1987]). They in turn are correlated with ordered maps of orientation and of direction of motion selectivity. All these findings are essentially the two-dimensional analogs of the auditory cortical maps. Thus, the orientation map is basically a gradient map in exactly the same sense used earlier in the AI map. This is because the orientation of an edge or a bar is basically a measure of its gradients in two directions. Similarly, the correlation between the orientation of a receptive field and its direction of motion selectivity is analogous to that found between a receptive field symmetry and its selectivity to FM sweeps. Finally, the spectral shape factor S also generalizes to two dimensions to reflect narrow elongated bars, edges of half-planes, and planer surfaces. All these features have long been recognized as physiologically and perceptually important attributes in form vision (Zeki and Shipp [1988]).

The similarity of cortical auditory and visual principles of processing is consistent with conclusions of studies into the generation of the neocortex and its subsequent parcellation into distinct areas (O'Leary [1989]). Particularly relevant for our arguments are experiments in which visual inputs from the optic nerve are induced in newborn ferrets to project to the audi-

tory cortex through the medial geniculate body (Sur, Garraghty and Roe [1988]). In the adult brains of such animals, AI cells can be shown to possess many of the response characteristics typical of the normal primary visual cortex, such as orientation selectivity. These and other manipulations, such as the transplanting of pieces of fetal neocortex to different positions, have all pointed to the homogeneity of the neocortex at its early stages of development, and the importance of subsequent influences, especially through afferent inputs, in differentiating the adult neocortical areas (O'Leary [1989]).

Acknowledgements

We are grateful for the help of Mr. Preetham Gopalaswamy and Dr. Radu Jasinschi throughout the course of this work. The Whitaker Foundation and the Naval Research Laboratory supported the physiological experiments. The National Science Foundation and the Air Force Office for Scientific Research provided grants for the theoretical and computational aspects of the work.

Figure Legends:

Figure 1

Schematic of ferret primary auditory cortex (AI). The tonotopic axis runs in the mediolateral direction with low frequencies laterally. 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 frequencies is 0.5-1 mm.

Figure 2

Schematic of basic findings of physiological experiments (Shamma, Fleshman and Wiser [1990]). Each box represents the AI, with ordinate representing the tonotopic axis from low (lf) to high (hf) frequencies, and the abscissa representing the (R)ostro-(C)audal axis. Dotted line represents a particular isofrequency plane at best frequency (BF) f_o.

Fig.2A: Schematic of receptive field organization in ferret AI. Each small plot represents the receptive field of a cell located along the indicated isofrequency plane, i.e., with $BF = f_o$. The excitatory and inhibitory responses as a function of frequency are symbolized by the clear and

shaded regions, respectively. Cells near the center of AI have symmetric receptive fields, with narrow inhibition flanking a narrow excitatory tuning curve. Towards the edges, receptive fields become more asymmetric and broader. Inhibition from high frequencies becomes relatively predominant caudally and weak rostrally, while low frequency inhibition exhibits the opposite trend.

Fig.2B: Schematic of AI responses to spectrally shaped noise stimuli. Each small plot represents the spectrum of the most effective stimulus for cells at that location along the isofrequency plane. Near the center of AI, cells best respond to narrow-band noise centered around f_o. In the caudal region, 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.

Fig.2C: Schematic of AI responses to frequency-modulated (FM) tones. Each arrow represents the most effective direction of the FM sweep for cells at that location along the isofrequency plane. Near the center, cells are least selective, responding well to both directions. Caudally, cells become more responsive to sweeps from low to high frequencies, i.e. arriving from the receptive field region with least inhibition. The opposite occurs in the rostral region.

Figure 3

Typical responses of cells with three types of side-band inhibition.

Middle Raster and Plot: Raster of responses to a *two-tone* stimulus. The cell's BF is 8.5 kHz. $T1$ starts at 100 ms into the sweep; its intensity is indicated to the right of the raster (65 dB SPL); its frequency spans ± 2 octaves around the BF in 1/4 octave steps. The second tone, $T2$, is delayed by 30 ms relative to $T1$; frequency = BF; intensity = 60 dB SPL. Ten repetitions are made at each $T1$ frequency.

The phasic responses of the cell to the two tones are segregated because of the inter-tone delay. $T1$ responses start approximately 16 ms following the onset of the tone (i.e., at 116 ms), and are restricted to a narrow range of frequencies around the BF (8 kHz). $T2$ responses start at 144 ms. They are vigorous when $T1$ is not near BF and are suppressed otherwise. Lateral inhibition is evidenced by the near absence of responses to either tone at the frequencies marked by the two arrows. *Plot below the raster* shows the total spike count as a function of

$T1$ frequency; the window is 80 ms long (100–180 ms).

Left Raster and Plot Typical responses of a cell with asymmetric inhibition from below the BF. The arrow marks the frequency of the one-sided inhibition. *Plot below the raster* shows the corresponding total spike count curve.

Right Raster and Plot Typical responses of a cell with asymmetric inhibition from above the BF. The arrow marks the frequency of the one-sided inhibition. *Plot below the raster* shows the corresponding total spike count curve.

Figure 4

Averaged responses of the three types of receptive fields measured in AI. The raster for each group is formed by collapsing comparable responses from many cells across different animals. The rasters span ± 2 octaves around the BF, at 1/4 octave steps. The *middle row of plots below the rasters* represent the average spike count curve for the each raster computed in the window 100–180 ms. Note that frequency axis is now labeled relative to a nominal BF. The ordinate represents the average number of spikes per cell. The *lower row of plots* represent the spike count curves computed from rasters that spanned only ± 1 octave around the BF, not necessarily from the same cells as in upper plots and rasters.

Figure 5

Superimposed plots illustrating the changes between the physiologically measured receptive field curves at the center and at the rostral (left) and caudal (right) edges of AI. In each plot, the *solid* curve represents the center receptive field (also shown in bottom of Fig.4), while dashed curves represent the rostral (left plot) and caudal (right plot) receptive fields. The dashed axis represents the baseline responses measured at the ends of the curves. Vertical alignment of the curves is based on matching their baselines.

Figure 6

Construction of the model receptive fields along the isofrequency plane f_o using excitatory and inhibitory gaussian functions. Upper row illustrates the variations in the width of the excitatory gaussian away from the center of AI ($x = 0$) towards the caudal region (shown at $x = 10$, and at the maximally caudal location, $x_m = 20$), and the rostral region ($x = -10$

and $x_m = -20$). The corresponding inhibitory gauusians at the same locations (second row) maintain their widths, but shift their means relative to f_o. The composite curves (last row) display the three primary features of the physiological receptive fields.

Figure 7

Superimposed curves illustrating the match between the physiological receptive fields (dashed curves) and the corresponding gaussian models (solid curves). The ordinate scales are the same as in Fig.4. The dashed axes represents the baseline of the model receptive fields. In each plot, the match is achieved by fixing the position of the two curves relative to the BF, and then sliding the model curve along the vertical axis until the central portions of the receptive fields approximately coincide.

Figure 8

Distribution of activity in the cortical model due only to the first term in Eq. 2 in the text. The box represents the surface of AI, with the (R)ostro-(C)audal axis running horizontally (from $x = -20$ to $x = 20$), and the tonotopic axis vertically. The central contour of AI ($x = 0$) is marked by the dotted line. The synthetic input pattern is plotted to the right of the map against the AI tonotopic axis. In all such plots, frequencies increase from bottom to top, as indicated schematically in Figs.2 and 3.

Figure 9

Schematics clarifying the linear contributions of the 2^{nd} term in Eq.2 in the text to the distribution of cortical activity. *The plots to the left* illustrate three examples of spectral input peaks centered at f_o (upper row), their derivatives (middle row), and their smoothed derivatives (lower row). The dotted line marks the location of the peak *at* f_o for all curves. Note that the value of the smoothed gradient *at* f_o changes depending on the symmetry of the input spectral peak, being zero for a symmetric peak (center), and positive (left) or negative (right) otherwise.

On the right, the plot illustrates the shape of $o(x)\bar{I}_t^{(1)}(f_o)$ as a function of x, for a positive and a negative value of the smoothed gradient $\bar{I}_t^{(1)}(f_o)$.

Figure 10

Schematics clarifying the contributions of the 3^{rd} term in Eq.2 in the text to the distribution of cortical activity. *The plots to the left* illustrate an example of a spectral input peak centered at f_o (upper), its inverted second-derivative (middle), and its smoothed derivative (lower) using the excitatory and the inhibitory (darker curve) gaussian windows. The dotted line marks the location of the peak *at* f_o for all curves. Note that the value of the smoothed second-derivative *at* f_o is larger in the case of the narrower excitatory gaussian.

On the right the curves illustrate the shapes of the even functions that constitute the 3^{rd} term in Eq.2 in the text ($e_1(x)\bar{I}_e^{(2)}$- and $e_2(x)\bar{I}_i^{(2)}$-), and the composite curve (the darker curve). Note the resultant curve is concave because the more curved component ($e_1(x)\bar{I}_e^{(2)}$-) is concave.

Figure 11

Correspondence between the cortical distributions of activity computed from the approximate expressions (Eq.2) and the original receptive fields (Eq.1), using a synthetic input pattern (left pair) and a natural speech token (right pair). For each pair, the approximate outputs are shown to the right. Details of the cortical plot are as in Fig.8. The frequency axis for the right pair of plots is derived from a model of the cochlea and lateral inhibitory networks that were used to compute the pattern. These models are described in detail in (Shamma [1989]).

Figure 12

Interpretations of the Spectral Shape Factor.

Fig.12A: Schematic plots of a cortical receptive field spectral peak at f_o (upper), a spectral peak at f_o flanked to the (L)eft and (R)ight by relatively constant slopes (second), and the 1^{st} and inverted second-derivatives of the input pattern (third and last).

Fig.12B: Schematic plots illustrating the characteristic values of the ratio $\frac{L+R}{L-R}$ for three different kinds of spectral shapes: Symmetric spectral peaks, asymmetric peaks, and edges.

Fig.12C: Progressive smoothing of the input spectral pattern changes the *local* peak around f_o from a symmetric peak (upper) to an asymmetric peak (lower) as the interference from the second peak increases. While upper curve shows all details of the pattern, lower curve reflects approximately its envelope.

Figure 13

Basic features of the cortical model outputs.

Fig.13A: Cortical model outputs using a synthetic pattern with three peaks of different symmetries. Details of plot axes are as in Fig.8.

Fig.13B-D: Cortical model outputs for a pattern with three symmetric peaks, with gradually decreasing spacing. Details of plot axes are as in Fig.8.

Figure 14

Outputs of the cortical model for various speech phonemes.

Fig.14A: Cortical model outputs for the phoneme /d/. Details of plot axes are as in Fig.8.

Fig.14B: Cortical model outputs for the phoneme /g/. Details of plot axes are as in Fig.8.

Fig.14C: Cortical model outputs for the phoneme /aa/. Details of plot axes are as in Fig.8.

Fig.14D: Cortical model outputs for the phoneme /u/. Details of plot axes are as in Fig.8.

References

P. ASSMANN AND Q. SUMMERFIELD, *Modelling the perception of concurrent vowels: Vowels with the same fundamental frequency* , J. Acoust. Soc. Am. , 85(1988), p. 327 .

T. BOURK, J. MIELCARZ AND B. NORRIS, *Tontopic organization of the anteroventral cochlear nucleus of the cat.*, Hearing Res., 4(1981), pp. 215–241.

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

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

T. IMIG AND H. ADRIAN, *Binaural columns in the primary field (AI) of cat auditory cortex.* Brain Res. , 138(1977), pp. 241–257 .

J. JONES AND L. PALMER, *The Two-dimensional spatial structure of simple receptive fields in cat striate cortex* , J. Neurophysiol. , 58(1987), pp. 1187–1211 .

W. D. KEIDEL AND W. D. NEFF, *Handbook of Sensory Physiology*, Berlin(1975).

D. KLATT, *Prediction of perceived phonetic distance from critical-band spectra: A first step* , Proc. ICASSP , 82(1982), p. 1278 .

M. LEVINE, *Vision in Man and Machine*, McGraw-Hill, New York, 1985.

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

J. MENDELSON, C. SCHREINER, K. GRASSE AND M. SUTTER, *Spatial distribution of responses to FM sweeps in cat primary auditory cortex*, Proc. 11th A.R.O. meeting, 1988.

M. MERZENICH, P. KNIGHT AND G. ROTH, *Representation of cochlea within primary auditory cortex in the cat*, J. Neurophysiol., 28 (1975), pp. 231–249 .

J. MIDDLEBROOKS, P. DYKES AND M. MERZENICH, *Binaural response-specific bands in primary auditory cortex (AI) of the cat: Topographical organization orthogonal to isofrequency contours*, Brain Res. , 181 (1980), pp. 31–48 .

W. NEFF, W. DIAMOND AND J. CASSEDAY, *Behavioral studies of auditory discrimination: Central nervous system.* , in Handbook of Sensory Physiology, W. Keidel and W. Neff, eds., 2 , Springer-Verlag, Berlin, 1975, pp. 307–400..

D. O'LEARY, *Do cortical areas emerge from a protocortex?*, Trends in Neuroscience, 12 (1989), pp. 400–406.

R. REALE AND T. IMIG, *Tonotopic organization of auditory cortex in the cat.* , J. Comp. Neurol. , 192 (1980), pp. 265–291 .

C. SCHREINER AND G. LANGNER, *Periodicity coding in the inferior colliculus of the cat. II. Topographical organization*, J. Neurophysiol., 60 (1988), pp. 1823–1840..

C. SCHREINER, J. MENDELSON, K. GRASSE AND M. SUTTER, *Spatial distribution of basic response properties in cat primary auditory cortex*, Proc. 11th A.R.O. meeting, 1988.

C. SCHREINER AND J. URBAS, *Representation of amplitude modulation in the auditory cortex of the cat. II. Comparison between fields.* , Hearing Res. , 32 (1988), pp. 49–64 .

S. SHAMMA, *Spatial and Temporal Processing in Central Auditory Networks*, in Methods in Neuronal Modelling, C. Koch and I. Segev, eds., MIT Press, Cambridge, 1989.

S. SHAMMA, J. FLESHMAN AND P. WISER, *Receptive Field Organization in Primary Auditory Cortex: Spectral Orientation Columns.*, Submitted to J. Neurophys., 1990.

M. SLANEY AND R. LYON, *Perceptual Pitch Detector*, Proc. ICASSP, 1990.

M. SUR, P. GARRAGHTY AND A. ROE, *Experimentally induced visual projections into auditory thalamus and cortex*, Science, 242 (1988), pp. 1437–1441 .

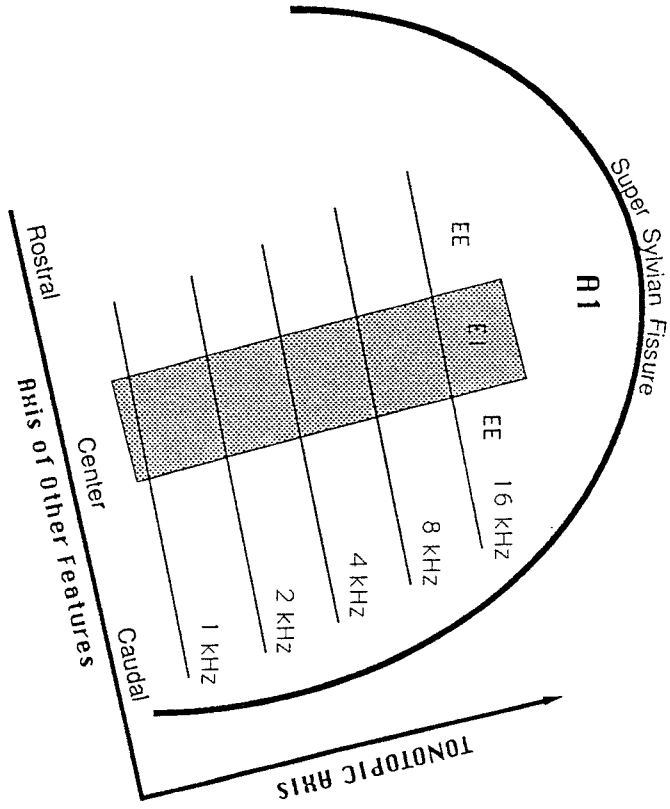
J. WENSTRUP, L. ROSS AND G. POLLAK, *Binaural response organization within a frequency band representation of the inferior colliculus: implications for sound localization* , J. Neuroscience, 6 (1986), pp. 962–973,.

I. WHITFIELD, *Auditory cortex and the pitch of complex tones*, J. Acoust. Soc. Am. , 67 (1980), pp. 644–647 .

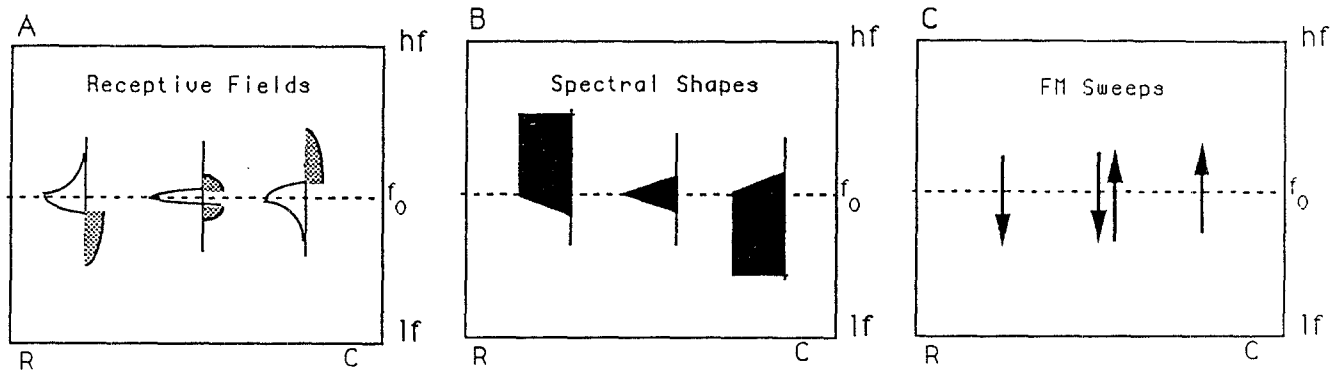
I. WHITFIELD AND E. EVANS, *Responses of auditory cortical neuron to stimuli of changing frequency*, J. Neurophysiol., 28 (1965), pp. 655–672.

P. WINTER AND H. FUNKENSTEIN, *The effect of species-specific vocalization on discharge of auditory cortical cells in the awake squirrel monkey.* , Exp. Brain Res. , 18 (1973), pp. 489–504 .

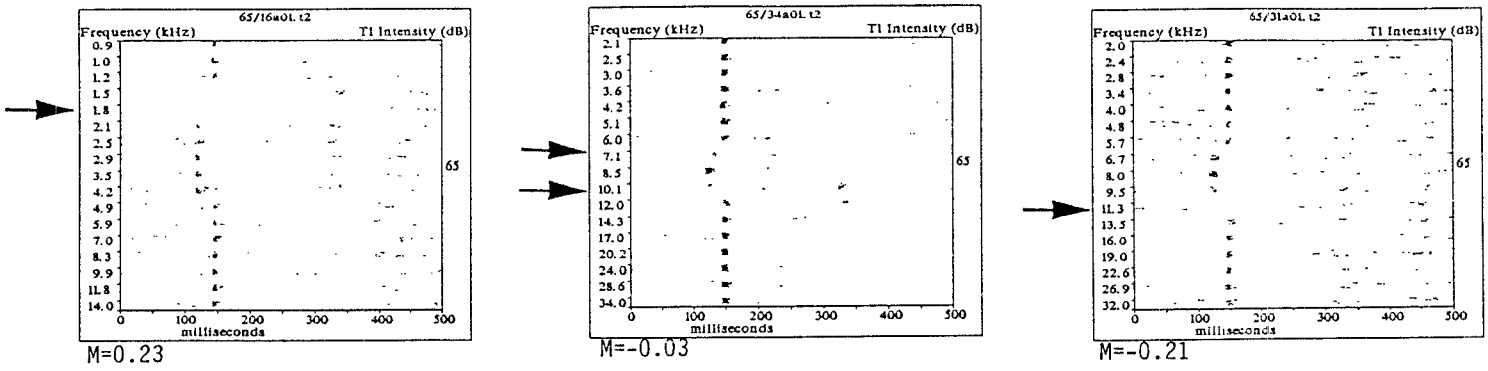
S. ZEKI AND S. SHIPP, *The functional logic of cortical connections*, Nature , 335 (1988), pp. 311–317 .

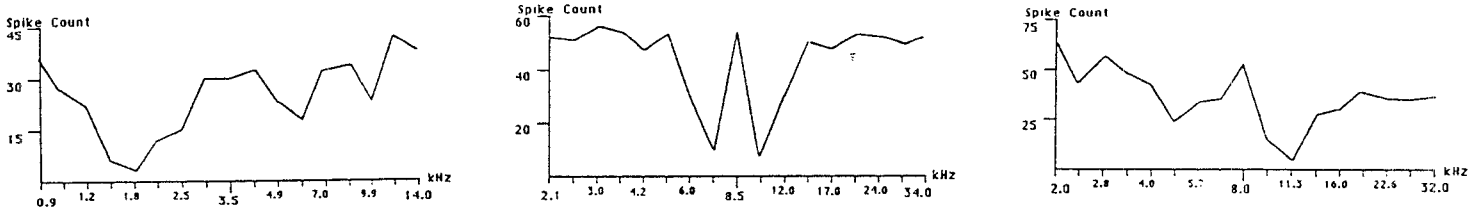


Shihab Shamma Fig 1



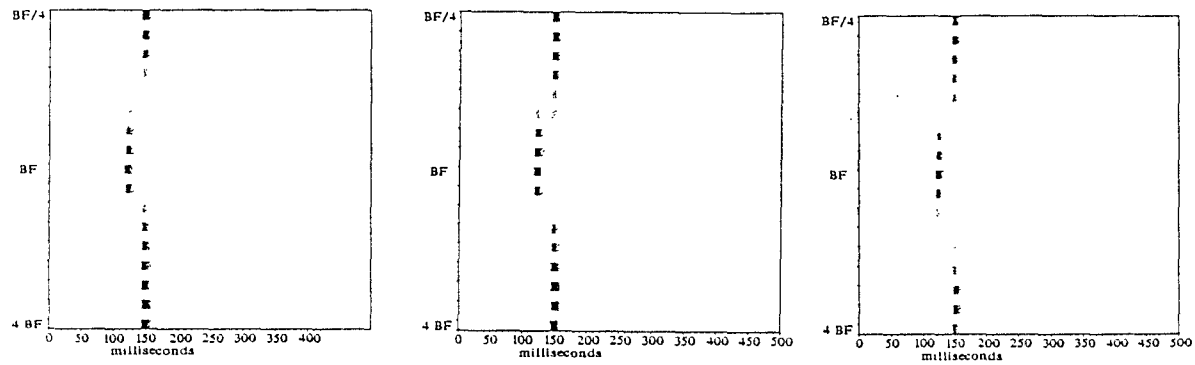
Shihab Shamma Fig 2

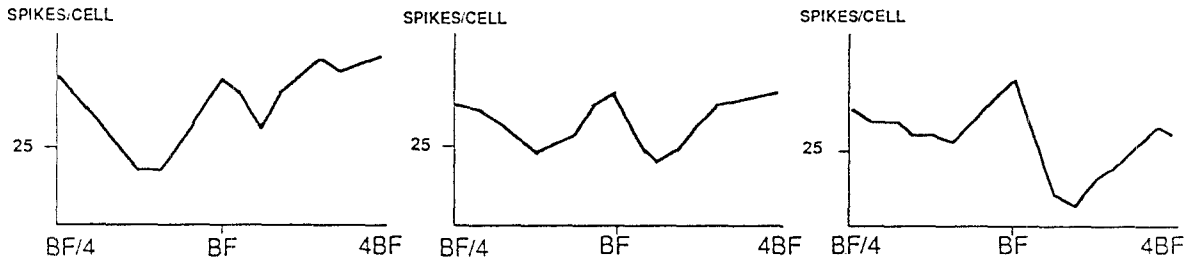


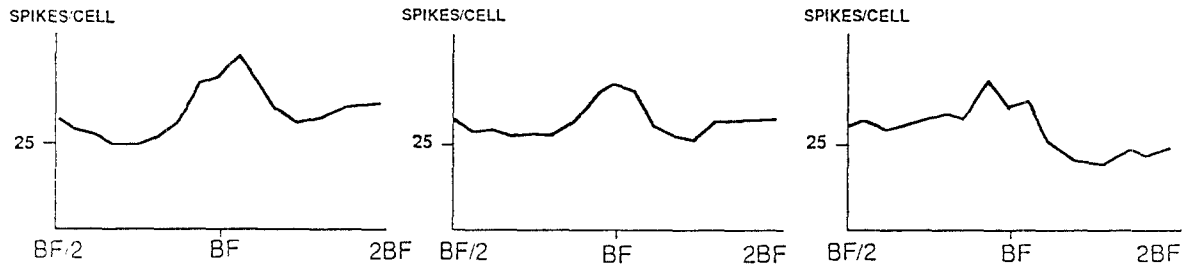


Shihab Shamma Fig 3

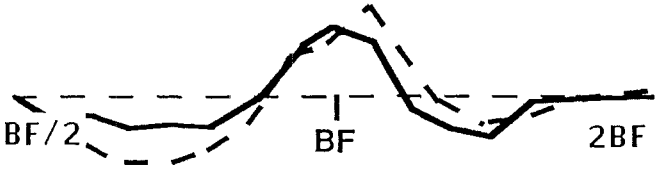
ROSTRAL CENTER CAUDAL



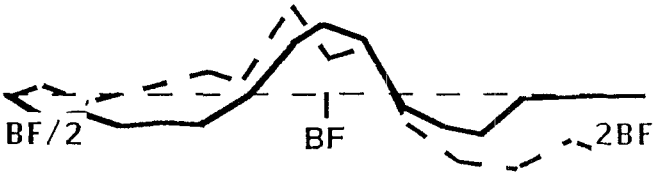




Shihab Shamma Fig 4

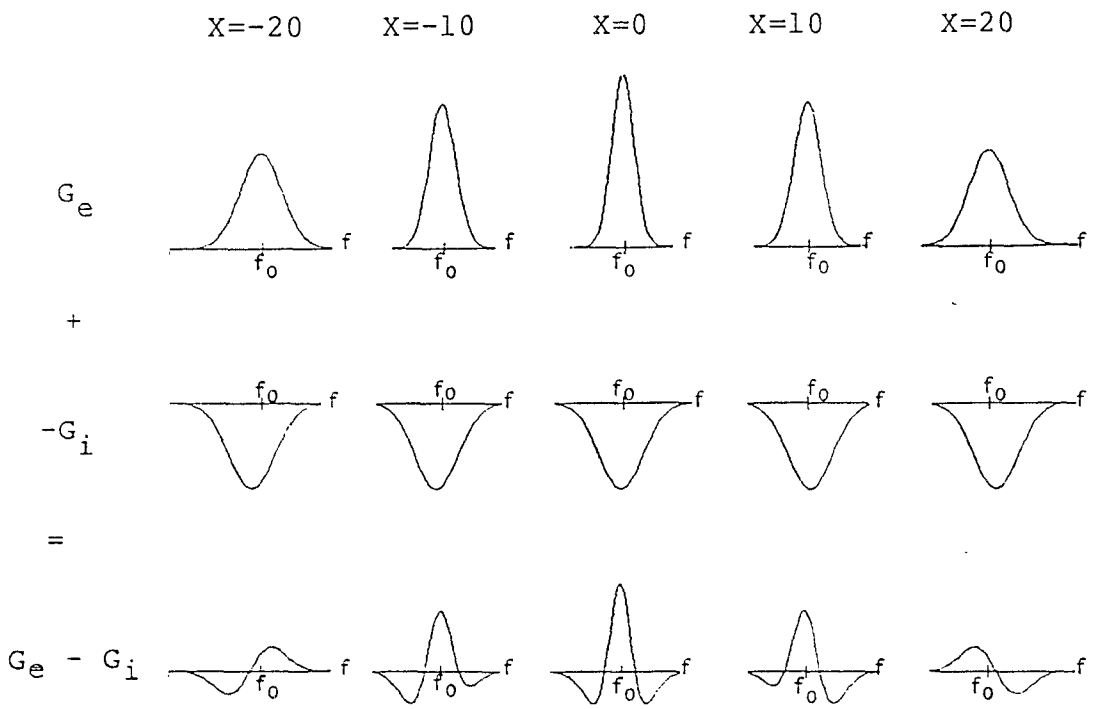


CENTER vs. ROSTRAL
RECEPTIVE FIELDS

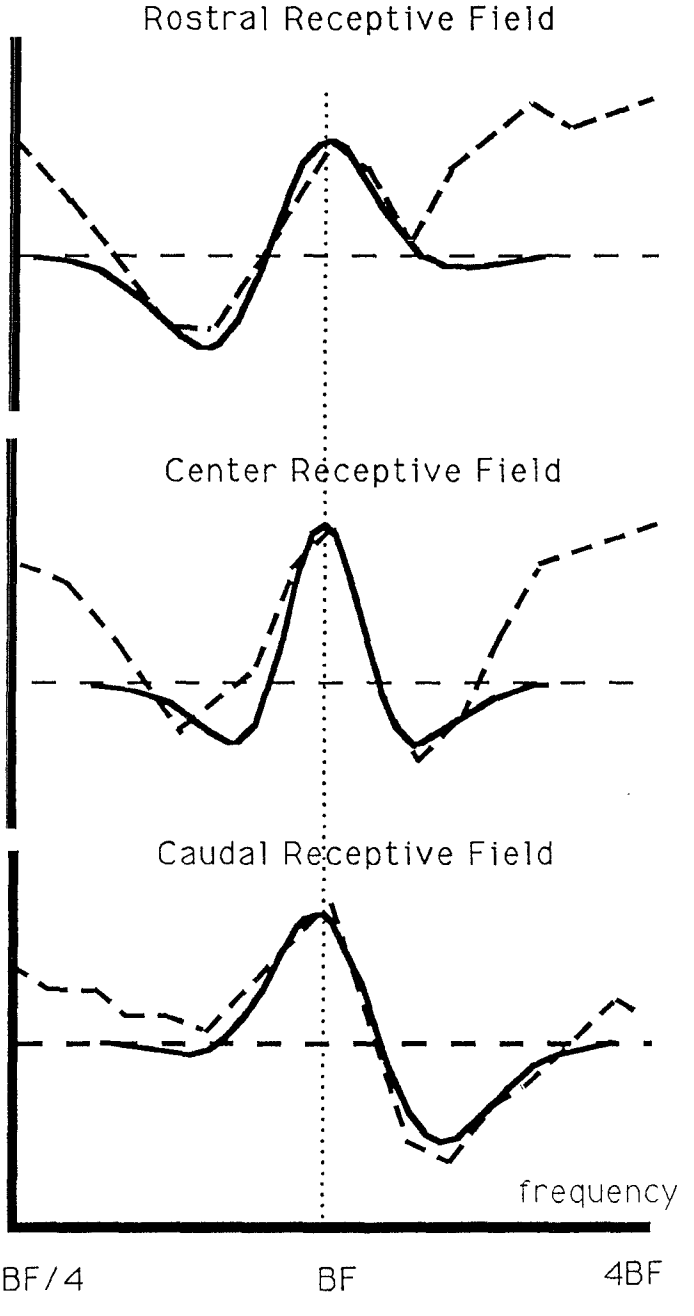


CENTER vs. CAUDAL
RECEPTIVE FIELDS

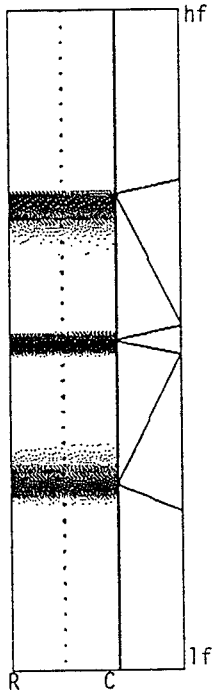
Shihab Shamma Fig 5



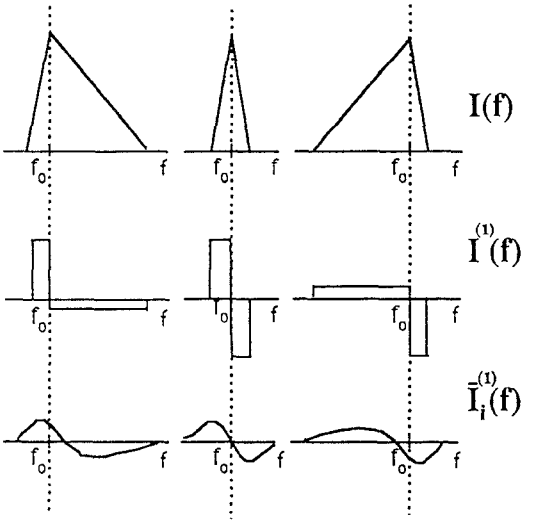
Shihab Shamma Fig 6

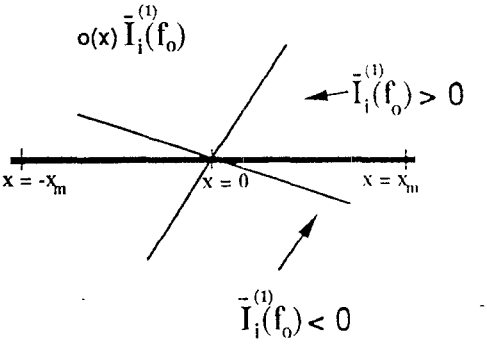


Shihab Shamma Fig 7

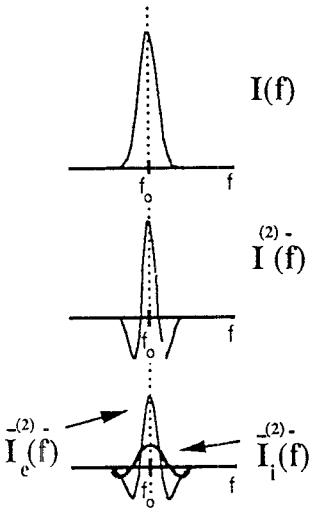


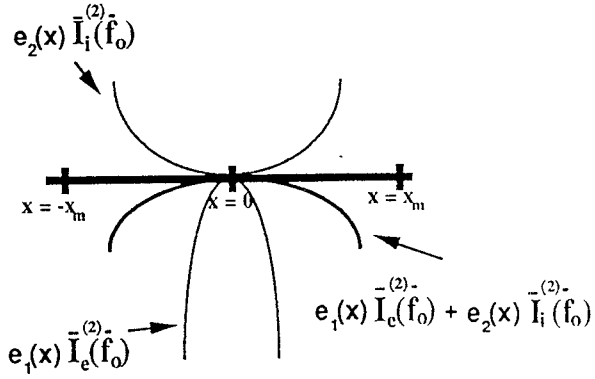
Shriab Chamber Fig 8



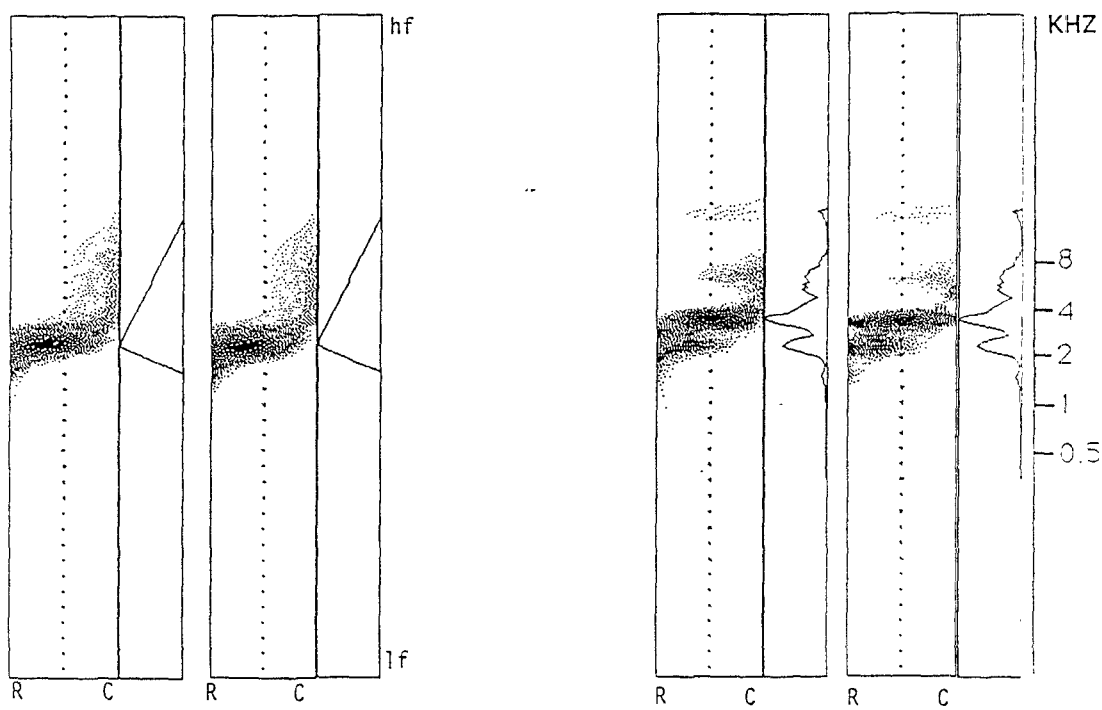


Shihab Shamma Fig 9

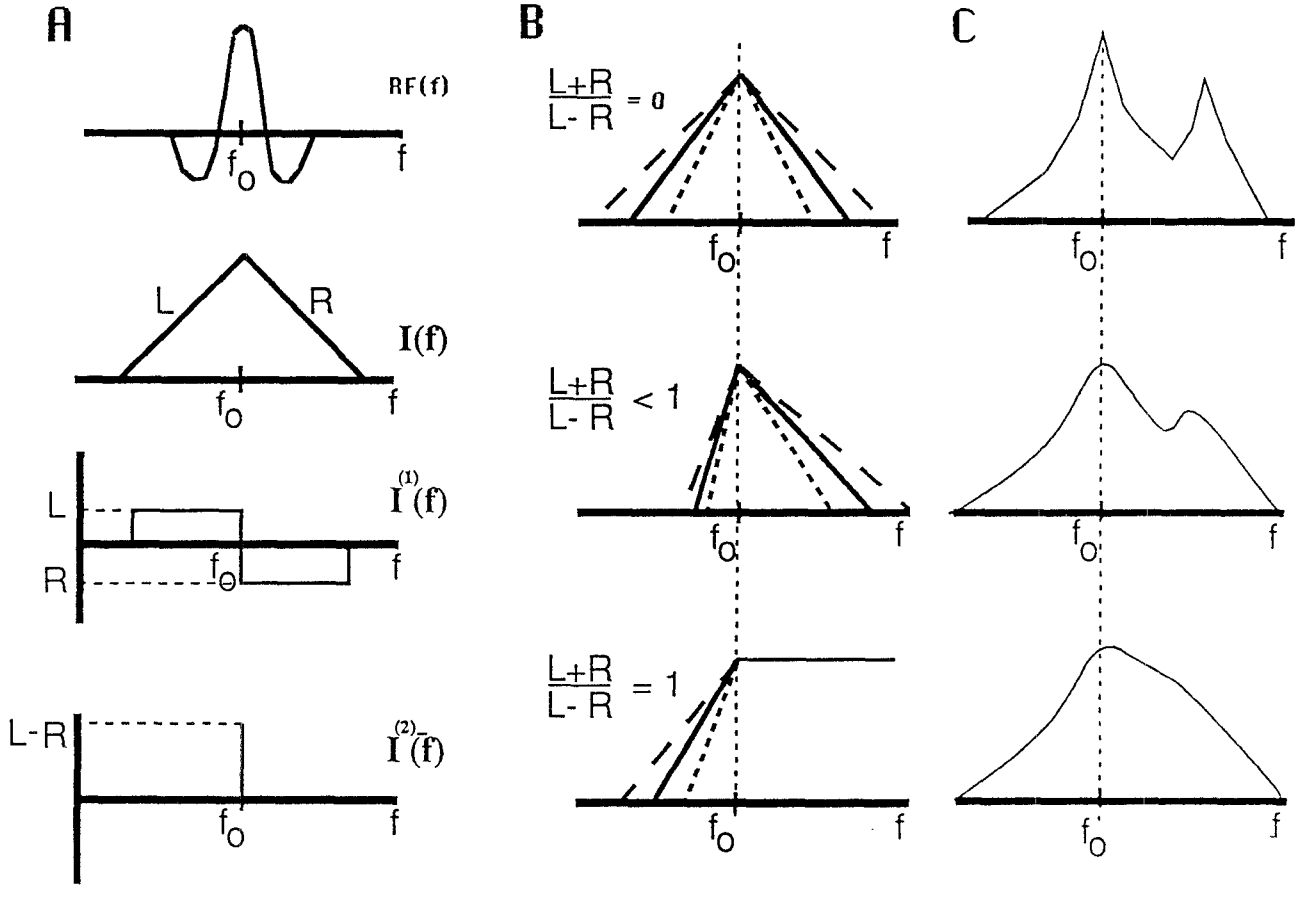




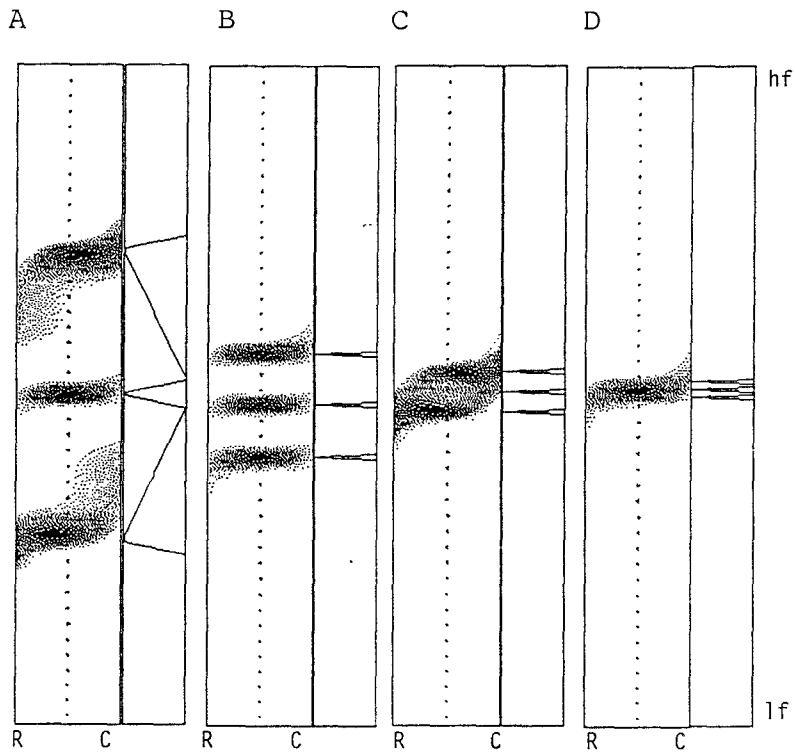
Shihab Shamma Fig 10



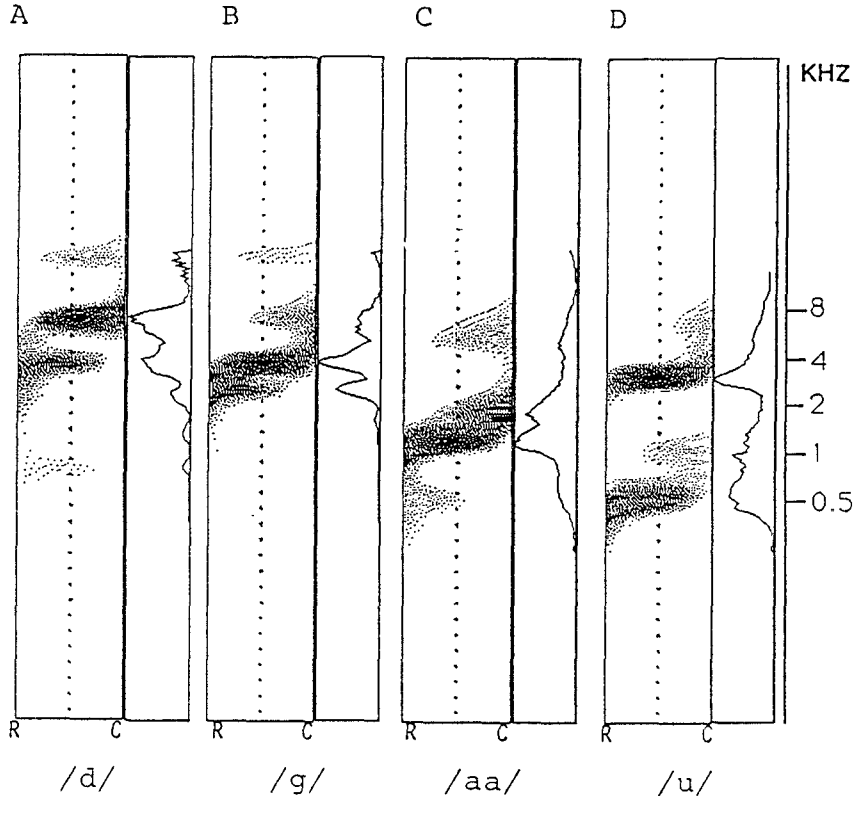
Shihab Shamma Fig 11



Shihab Shamma Fig. 12



Shihab Srinivas Fig 13



Shihab Shamma Fig 12A