*Auditory Neuroscience*, 1995, Vol 1, pp. 255–270
Reprints available directly from the publisher
Photocopying permitted by license only



# Ripple Analysis in Ferret Primary Auditory Cortex. II. Prediction of Unit Responses to Arbitrary Spectral Profiles

SHIHAB A. SHAMMA and HUIB VERSNEL

*Electrical Engineering Department and Institute for Systems Research, University of Maryland, College Park, Maryland 20742*

(*Received October 27, 1994, accepted February 22, 1995*)

**We examined whether responses in primary auditory cortex (AI) to arbitrary spectral profiles can be explained by the superposition of responses to the individual ripple components that make up the spectral pattern. For each unit, the ripple transfer function was first measured using ripple stimuli consisting of broadband complexes with sinusoidally modulated spectral envelopes (Shamma *et al.* 1995). Unit responses to various combinations of ripples were compared to those predicted from the superposition of responses according to the transfer function. Spectral profiles included combinations of 2 to 5 ripples of equal amplitudes and random phases, and vowel-like profiles composed of 10 ripples with various amplitudes and phases. The results demonstrate that predicted and measured responses are reasonably well matched and, hence, support the notion that AI analyzes the acoustic spectrum in a substantially linear manner.**

*Key words:* Primary auditory cortex, spectral ripples, spectral profiles, superposition, ferret

THE ACOUSTIC SPECTRAL PROFILE is a primary cue in the perception of timbre (Plomp, 1976). A fundamental goal in auditory cortical physiology has been to understand how this profile is represented in the firing rate of cortical cells, or, equivalently, how might one predict the responses of a single unit to arbitrary spectral profiles. In primary auditory cortex (AI), the most commonly used descriptor of unit responses has been the frequency response area (Phillips *et al.*, 1991). As conventionally defined and measured, it provides information about the best (or characteristic) frequency (BF), and surrounding excitatory and inhibitory influences on the cell. Although these measures have been useful as a qualitative guide to the responses expected of a single unit to tonal and other narrowband stimuli, they are not suitable for precise quantitative predictions of responses to arbitrary spectral profiles.

More appropriate response area measures could be derived from the responses of AI cells to broadband rippled spectra, that is, spectra with sinusoidal envelopes (Schreiner and Calhoun, 1995; Shamma *et al.*, 1995). Specifically, each unit could be characterized by a so-called "ripple tranfer function" that reflects the magnitude and phase of its response to different ripple frequencies. Most AI cells exhibit band-pass transfer functions that are tuned around a characteristic ripple frequency and phase. These latter two parameters are roughly correlated to the bandwidth and asymmetry of the response area (Shamma *et al.*, 1995). Based on this finding, it was concluded that most AI cells exhibit a linear component in their responses. Under the assumption of linearity, it is theoretically possible to predict the responses of a unit to any spectral profile by applying the "principle of superposition." Following this principle, the profile is decomposed into its constituent ripple components, and then the weighted contributions of each ripple component are summed according to the cell's ripple transfer function.

In this report, we examine directly the extent to which ripple superposition (and hence the linearity of the system) holds. Specifically, we shall compare the responses of AI cells to various combinations of ripples with those predicted from their ripple transfer functions.

## METHODS

### *Surgery and Animal Preparation*

The ferrets were anesthetized with sodium pentobarbital (40 mg/kg). Anesthesia was maintained throughout the experiment by continuous intravenous infusion of

Corresponding author: Mr. Shihab Shamma, Electrical Engineering Dept., Univ. of Maryland, College Park, M.D. 20742.

pentobarbital. The ectosylvian gyrus, which includes AI, was exposed by craniotomy and the dura was reflected. The contralateral ear canal was exposed and partly resected, and subsequently a cone-shaped speculum containing a Sony MDR-E464 miniature speaker was sutured to the meatal stump. For details on the surgery, see Shamma *et al.* (1993 in this issue).

### *Acoustic Stimuli*

For each cell, we measured a frequency response curve with up to 1/8 octave resolution at low intensity. The BF was determined from this response curve as the frequency that evoked the best response (thus, BF approximates the frequency of the lowest threshold). The rate-level function at BF was measured at a range from 35 to 85 dB sound pressure level (SPL) to determine the cell's response threshold and the nonmonotonicity. The criteria were 10% of maximum response and a decrease of 25% with increase of intensity, respectively.

All other stimuli used in these experiments were broadband complex sounds consisting of 101 tones that were equally spaced along the logarithmic frequency axis and spanning 4.32 octaves (such as 1 to 20 kHz or 0.5 to 10 kHz), as illustrated in Figure 1. The range was chosen so that the response area of the cell tested lay within the stimulus' spectrum. The spectral envelope of the complex was then modulated in one of two ways, either as a single sinusoid along the frequency axis on a linear or logarithmic amplitude scale (Fig. 1A) or as a waveform representing the superposition of several sinusoids (Fig. 1B).

The overall level of the complex stimulus was defined by the level of a single frequency component, $L_1$ dB SPL in the flat complex. Thus, the overall level for a flat complex with 101 components (ripple amplitude $\Delta A$ at zero) was taken to be $L_1 + 10 \log(101) \approx L_1 + 20$ dB. The overall stimulus level was chosen on the basis of the threshold at BF, typically $L_1$ was set about 10 dB above threshold. High levels ($L_1 > 65$ dB SPL) were avoided to ensure the linearity of our acoustic delivery system. The amplitude of a single ripple was defined as the maximum percentage or logarithmic change in the component amplitudes. Ripple amplitudes were at 90 to 100%, or 10 dB, modulation. In a few cases, different ripple amplitudes and stimulus levels were tried.

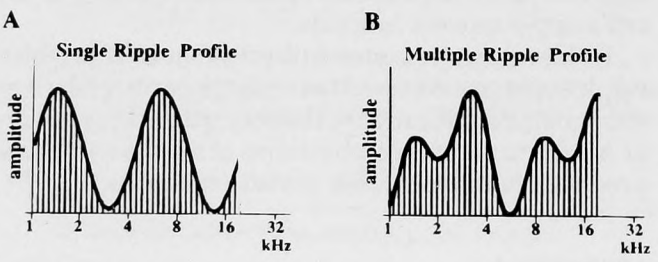



*FIGURE 1* **(A) Schematic of a stimulus spectrum composed of a single ripple at 0.5 cycle/octave and 0° phase.** Its amplitude is defined as a 100% linear amplitude modulation. (B) Schematic of a stimulus spectrum composed of two ripples at 0.4 and 0.8 cycle/octave and phases −105° and −41°, respectively. The amplitude of the ripple complex is defined as a 100% linear amplitude modulation.

The ripple frequency ($\Omega$) is presented in units of cycles/octave against the logarithmic frequency axis. The ripple phase ($\Phi$) is presented in radians (or degrees) relative to a sine wave starting at the left edge (low-frequency edge) of the complex (Fig. 1A). To measure the ripple transfer function of a cell, a series of tests were carried out using rippled spectra with a range of ripple frequencies $\Omega$ (usually 0 to 2 cycles/octave with different resolutions) and ripple phases $\Phi$ (from 0 to $7\pi/4$ in $\pi/4$ steps). Each stimulus was typically repeated 20 times.

A multiple-ripple stimulus typically consisted of 2 to 5 ripple components. The relative amplitude and phase of each ripple was first specified. The compound waveform due to the superposition of all ripples was then generated and used to shape the envelope of the spectrum as before. The spectral range, overall level, and ripple amplitude of the compound ripple stimuli were set as in the single ripples.

The complex stimulus bursts had 7 ms rise and fall time and 50 ms duration. They were computer synthesized, gated, and then fed through a common equalizer into the earphone. Calibration of the sound delivery system (up to 20 kHz) was performed in situ using a 1/8 inch Brüel & Kjaer probe microphone (type 4170). The microphone was inserted into the ear canal through the wall of the speculum to within 5 mm of the tympanic membrane. The speculum and microphone setup resembles closely that suggested by Evans (1979).

### *Recordings*

Action potentials from single units were recorded using glass-insulated tungsten microelectrodes with 5 to 6 MΩ tip impedances. Neural signals were led through a window discriminator and the time of spike occurrence relative to stimulus delivery was stored using a Hewlett-Packard 9000/800 series minicomputer. The computer also controlled stimulus delivery, and created various raster displays of the responses.

In each animal, electrode penetrations were made orthogonal to the cortical surface. In each penetration, cells were typically isolated at depths of 350 to 600 μm corresponding to cortical layers III and IV (Shamma et al., 1993).

### *Data Analysis for Single Ripple Stimuli*

Figure 2 illustrates the display and initial analysis applied to the data. Details of these procedures are described in Shamma *et al.*, 1995. Here the cell was tested over ripple frequencies 0 to 2 cycles/octave in steps of 0.4 cycles/octave. For each ripple, the responses to a full cycle of the ripple (that is, $2\pi$ phase change) was measured at eight steps. The spike counts at each phase step

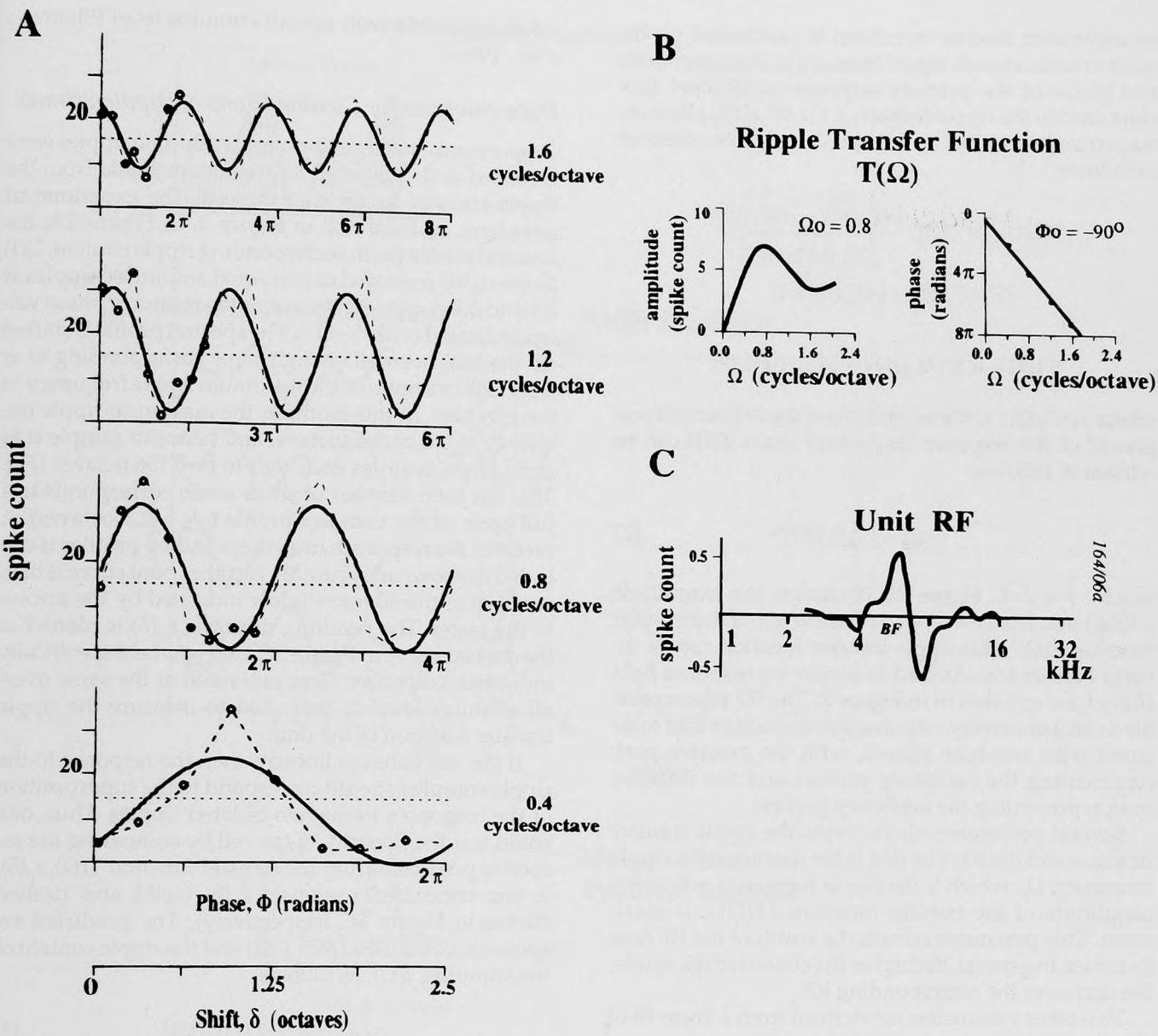



*FIGURE 2* **(*A*) Measured and fitted responses to single ripple profiles at various ripple frequencies.** In each plot, the response spike count to a ripple is measured at various phases of the ripple (eight $\pi/4$ steps per cycle) as indicated by the circles. The solid curve is the best sinusoidal fit to the data. For the 0.4 cycle/octave ripple, a full cycle of the response is equivalent to a 2.5 octave shift (or translation) of the stimulus profile, as indicated by the two axes at the bottom. For ripples 0.8 to 1.6 cycles/octave, the full cycle corresponds to progressively smaller shifts of the profiles. So in the plots, the periodic response curves are simply repeated to indicate what they would look like if the full 2.5 octave shift had been applied. The dotted baseline is the spike count obtained for the flat-spectrum stimulus. (*B*) The ripple transfer function $T(\Omega)$. The plot to the left represents the weighted amplitude of the fitted sinusoids (as in *A*) as a function of ripple frequency $\Omega$. $\Omega_o$ is the ripple frequency with the maximum response amplitude. The plot to the right represents the phases of the fitted sinusoids as a function of ripple frequency. The characteristic phase, $\Phi_o$ is the intercept of the linear fit to the data. (*C*) The response field (RF) of the unit. It is computed from the inverse Fourier transform of the ripple transfer function.

were made over a 60 ms time window starting shortly (10 ms) after the onset of the stimulus. These counts are indicated by the small circles, which are connected by the dashed lines, in the plots of Figure 2A. The baseline at each ripple frequency (represented by the dotted horizontal line) was set equal to the spike count obtained from the flat spectrum ($\Omega = 0$).

The axis at the bottom, labeled as $\delta$ (octaves), indicates the equivalent amount of shift each ripple pattern undergoes at each phase step. For instance, for a 0.4 cycles/octave ripple, response measurements over a full cycle are equivalent to shifting the spectral pattern by 2.5 octaves along the logarithmic frequency axis. The same phase steps for a 0.8 cycles/octave pattern are equivalent to shifting it by half as much (1.25 octaves). To estimate the ripple transfer function ($T(\Omega)$) of the cell,

an eightpoint Fourier transform is performed on the spike counts at each ripple frequency. The magnitude and phase of the primary response component synchronized to the ripple frequency Ω ($AC_1(\Omega)$) is then extracted and weighted by the *rms* value of the response as follows:

$$T(\Omega) = AC_1(\Omega) \cdot \frac{|AC_1(\Omega)| - |AC_1(0)|}{\sqrt{\Sigma_{i=1}^{4} |AC_i(\Omega)|^2}}$$
$$\text{if } |AC_1(\Omega)| - |AC_1(0)| \geq 0 \qquad (1)$$
$$T(\Omega) = 0 \text{ if } |AC_1(\Omega)| - |AC_1(0)| < 0$$

where $|AC_i(\Omega)|$ is the magnitude of the $i^{th}$ Fourier component of the response. In general terms $T(\Omega)$ can be written as follows:

$$T(\Omega) = |T(\Omega)| e^{j\phi(\Omega)} \qquad (2)$$

where $j = \sqrt{-1}$. Figure 2B illustrates the magnitude $|T(\Omega)|$ and the unwrapped phase $\Phi(\Omega)$ of the transfer function $T(\Omega)$. This ripple transfer function can be inverse Fourier transformed to obtain the response field (RF) of the cell shown in Figure 2. The RF is comparable to an isointensity response curve, such as that measured with two-tone stimuli, with the positive peak representing the excitatory portion and the negative peak representing the inhibitory portion.

Several parameters characterize the ripple transfer function and the RF. The first is the characteristic ripple frequency, $\Omega_o$, which is the ripple frequency where the magnitude of the transfer function, $|T(\Omega)|$, is maximum. This parameter reflects the width of the RF near its center. In general, the higher the characteristic ripple, the narrower the corresponding RF.

Two other parameters are derived from a linear fit of the phase function according to;

$$\Phi(\Omega) = x_o\Omega + \Phi_o \qquad (3)$$

where $x_o$ is the slope of the line and $\Phi_o$ is its intercept. The parameter $x_o$ reflects the location (in octaves) of the RF relative to the left edge of the ripple. The parameter $\Phi_o$ (called the characteristic phase) roughly indicates the asymmetry of the RF about its center. For instance, the RF is symmetrical for $\Phi_o = 0$, and strongly asymmetrical for $\Phi_o = \pm 90°$ (as in Fig. 2B). Another response parameter is the location of the maximum of the RF along the tonotopic axis. This has been shown to correspond well to the tonal BF of the cell and hence will be labeled as such in this article. The RF (or the ripple transfer function) was usually measured only at the stimulus level, which elicited a relatively strong response. This is justified by the fact that the RF remains relatively stable with overall stimulus level (Shamma *et al.*, 1995).

### *Data Analysis for Combinations of Ripple Stimuli*

Responses to spectra composed of multiple ripples were recorded and compared to predictions made from the ripple transfer function of the cell. The experimental paradigm is illustrated in Figure 3. In Figure 3A, the spectral profile [with corresponding ripple content, I(Ω) to the right] consisted of two equal amplitude ripples at 0.4 and 0.8 cycles/octave and at the arbitrary phase values indicated (–105°, –41°). The spectral profile is shifted (to the left) in small enough steps corresponding to at least eight samples of the maximum ripple frequency in the complex. In this example, the maximum ripple frequency is 0.8 cycles/octave, and hence to sample it in eight steps, requires each shift to be 0.156 octaves (Fig. 3B). The total number of shifts made corresponds to a full cycle of the complex profile ( $\frac{1}{0.4}$ = 2.5 octaves). A raster of the responses to all these shifted profiles is collected as shown in Figure 3B. A spike count curve is then made over the 60 ms window indicated by the arrows in the raster. The resulting response $r_m(\delta)$ is plotted as the dashed curve in Figure 3C. Except where specifically indicated, responses were measured at the same overall stimulus level as that used to measure the ripple transfer function of the unit.

If the cell behaves linearly, then the response to the ripple complex should correspond to the superposition of the responses to the two isolated ripples. Thus, one could test the linearity of the cell by comparing the response predicted from the transfer function $T(\Omega)$, $r_p(\delta)$, to the measured response, $r_m(\delta)$ (solid and dashed curves in Figure 3C, respectively). The predicted response is computed from $T(\Omega)$ and the ripple content of the stimulus, $I(\Omega)$, as follows:

$$r_p(\delta) = \mathfrak{F}^{-1}\{T(\Omega)\, I(\Omega)\} \qquad (4)$$

where $F^{-1}\{\cdot\}$ designates the inverse Fourier transform operation with respect to Ω, $I(\Omega)$ is the ripple content of the stimulus, and $T(\Omega)$ is the ripple transfer function. An equivalent way to compute the predicted response, which follows directly from Eq. 4, is to convolve the impulse response of the cell, or to cross-correlate the RF of the cell, $w(x)$, with the stimulus spectral profile, $p(x)$:

$$r_p(\delta) = \sum_x w(x + \delta)\, p(x) \qquad (5)$$

where $x$ is the logarithmic frequency axis (i.e., $x = \log_2$ frequency (kHz)). Therefore, the predicted response in Figure 3C is the sum of the curves fitted to the individual ripple responses (that is, the solid curves in the bottom two panels in Fig. 2A) except with each curve linearly amplitude scaled and phase shifted according

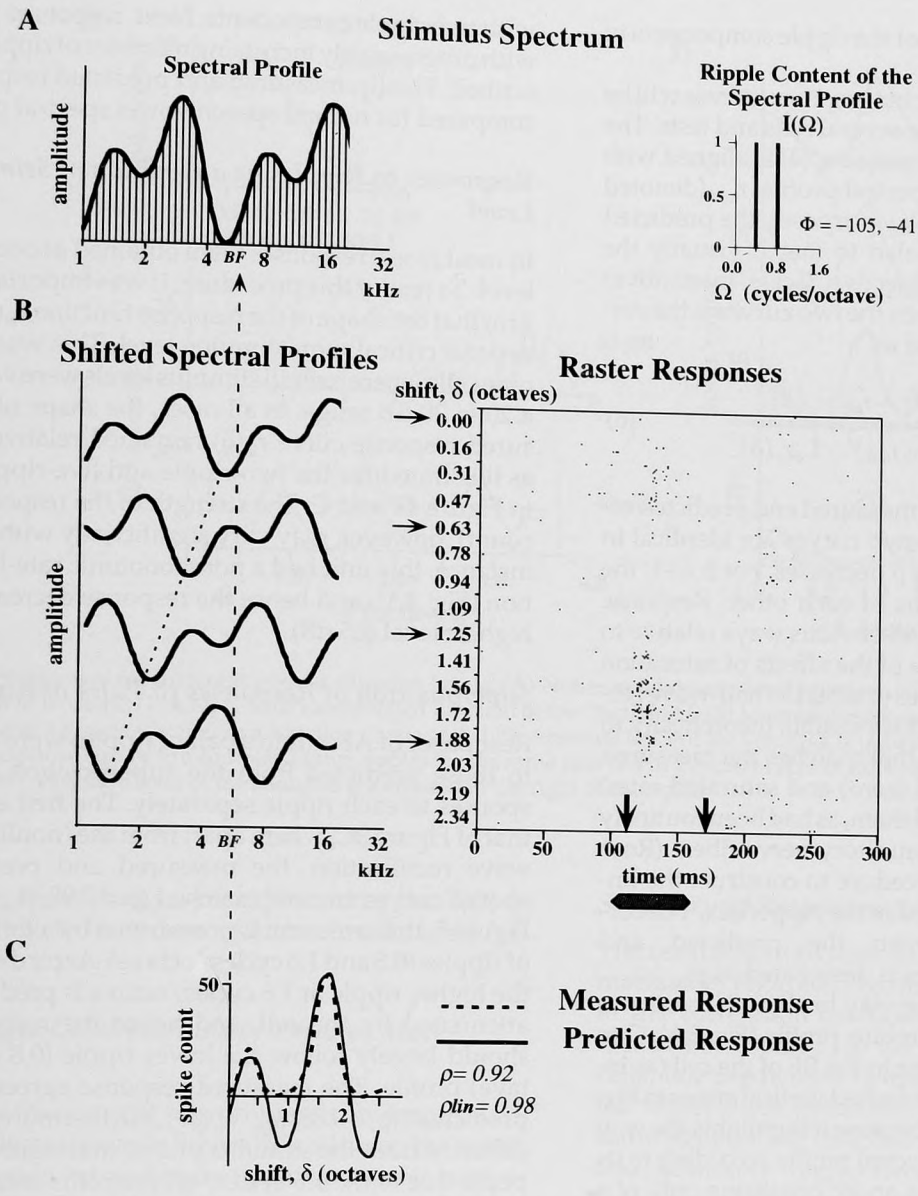



*FIGURE 3* **(*A*) The spectral profile of a stimulus (left plot) composed of two ripples.** The amplitude and phases of the two ripples are schematically illustrated in the right plot. (*B*) The spectral profile of the stimulus with increasing amount of shift (from top to bottom, as indicated by the dashed line). Note that the profile is periodic against the tonotopic axis with a period of 2.5 octaves. The underlying tones of the stimulus complex are omitted in these plots. The raster to the right illustrates the nature of the responses obtained as a function of profile shift. The profile is always shifted by a total amount equal to its period (2.5 octaves for this profile). The stimulus burst is indicated by the bar below the raster. The arrows define the window over which the response spike counts are made. (*C*) The response spike counts to different shifts are indicated by the dashed curve as a function of profile shift. The solid line is the response predicted from the ripple transfer function and the stimulus profile. The scale of the solid curve is in arbitrary linear units. The dotted horizontal line is the spike count of the flat spectral profile; it is used as the baseline for the predicted response curve, $r_f(\delta)$. The whole plot is aligned with the stimulus profile according the BF of the unit (determined from Fig. 2C). The correlation coefficients $\rho$ and $\rho_{lin}$ are indicated in the figure.

to the amplitude and phase of the ripple components in the stimuls, $I(\Omega)$.

The measured and predicted response curves will be illustrated as in Figure 3C for several cells and tests. The baseline of the predicted response $r_p(\delta)$ is aligned with the spike count of the flat spectral profile, $r_{m0}$ (denoted by the dotted line). For display purposes, the predicted curve is then arbitrarily scaled to match visually the measured response. An objective (scale insensitive) measure of the match between the two curves is the correlation coefficient $\rho$ defined as:

$$\rho = \frac{\Sigma_\delta (r_m(\delta) - r_{m0}) r_p(\delta)}{\sqrt{\Sigma_\delta (r_m(\delta) - r_{m0})^2 \cdot \Sigma_\delta r_p(\delta)^2}} \qquad (6)$$

where $r_m(\delta)$ and $r_p(\delta)$ are the measured and predicted response curves. If $\rho = 1$, the two curves are identical in shape; the match is worse as $\rho$ decreases. For $\rho = -1$, the curves are inverted versions of each other. Response curves were often distorted in obvious ways relative to the predicted curve because of the effects of saturation and rectification of firing rates (that is, the half-wave rectified response in Fig. 3C). It is possible theoretically to construct a response curve that matches the measured curve between the rectified (zero) and saturated rates, and extends linearly beyond them, as had been routinely done with responses from auditory nerve fibers (Rose *et al.*, 1967). A new fast procedure to construct the linearized response is described in the Appendix. The correlation coefficient between the predicted and constructed response curves is designated as $\rho_{lin}$.

Finally, there is no unique way to align the response curves (Fig. 3C) with the stimulus profile (Fig. 3A). One useful alignment is according to the BF of the cell (as indicated by the location of the dashed vertical arrow in Fig. 3). This alignment is useful because it highlights the way the cell distorts the input spectral profile according to its RF. For instance, if a cell has an RF consisting only of a narrow excitatory response area around the BF (that is, narrow relative to the details of the stimulus profile), then its responses would simply track the shape of the input profile as it is shifted past the BF. For such a cell, the response curve aligned with the BF would match the stimulus profile. If a cell's RF is asymmetrical or broad relative to the stimulus profile features (or equivalently, if some stimulus ripples are filtered out by $T(\Omega)$), the profile and the BF-aligned response curve would differ in shape.

## RESULTS

The data illustrated here were collected from a total of 51 single-unit recordings in five animals. All these units responded to tones and rippled stimuli. In this section, we first illustrate the dependence of the measured response functions $r_m(\delta)$ on absolute stimulus level, then compare measured and predicted responses to stimuli with two ripple components. Next, responses to stimuli with progressively increasing numbers of ripples are described. Finally, measured and predicted responses are compared for natural speech vowel spectral profiles.

### *Responses to Ripples as a Function of Stimulus Level*

In most cases, responses were obtained at one stimulus level. To justify this procedure, it was important to confirm that the shape of the response function $r_m(\delta)$ did not depend critically on stimulus level. This was tested in nine cells where overall stimulus levels were varied over a 20 to 30 dB range. In all cases, the shape of the measured response curve $r_m(\delta)$ remained relatively stable, as illustrated for the two-ripple and five-ripple stimuli in Figure 4B and C. The strength of the response (spike count), however, may vary significantly with level. For instance, this unit had a nonmonotonic rate-level function (Fig. 4A), and hence the response decreased at the highest level (65 dB).

### *Superposition of Responses to Pairs of Ripples*

Responses of AI cells to a pair of ripples were compared to those predicted from the superposition of the responses to each ripple separately. The first example is that of Figure 3C, where apart from the (nonlinear) half-wave rectification, the measured and predicted response curves are well matched ($\rho = 0.92$, $\rho_{lin} = 0.98$). In Figure 5, the same unit is now driven by a different pair of ripples (0.8 and 1.6 cycles/octave). According to $T(\Omega)$, the higher ripple at 1.6 cycles/octave is predicted to be attenuated by the cell, and hence the response $r_m(\delta)$ should largely follow the lower ripple (0.8 cycles/octave) profile. The measured response agrees with this prediction ($\rho = 0.88$, $\rho_{lin} = 0.88$). Furthermore, it is quite different from the stimulus profile that exhibits smaller peaks due to the 1.6 cycles/octave component.

These responses can also be interpreted as the cross-correlation of the RF (or convolution of the impulse response) with the spectral profile (see Methods). As such, the changes in $r_m(\delta)$ compared with the stimulus profile can be attributed to the shape of the RF. For instance, the absence of the smaller peaks (Fig. 5C) can be explained by the suppression induced by neighboring large peaks (to their right) via the inhibitory side-bands of the RF.

Examples from four other cells with different RFs are presented in Figure 6. In each case, the responses can be interpreted as the convolution of the RF with the stimulus profile. The responses in Figure 6B and D clearly illustrate the filtering effects of the RF (or the ripple transfer function) since $r_m(\delta)$ differs significantly from the corresponding stimulus spectrum. For example, in Figure 6B, the small peak in the stimulus spectral profile at 2.5 or 17 kHz evokes little corresponding response; in Figure 6D, the peaks of the spectral profile evoke responses with the opposite relative strength. Note that in

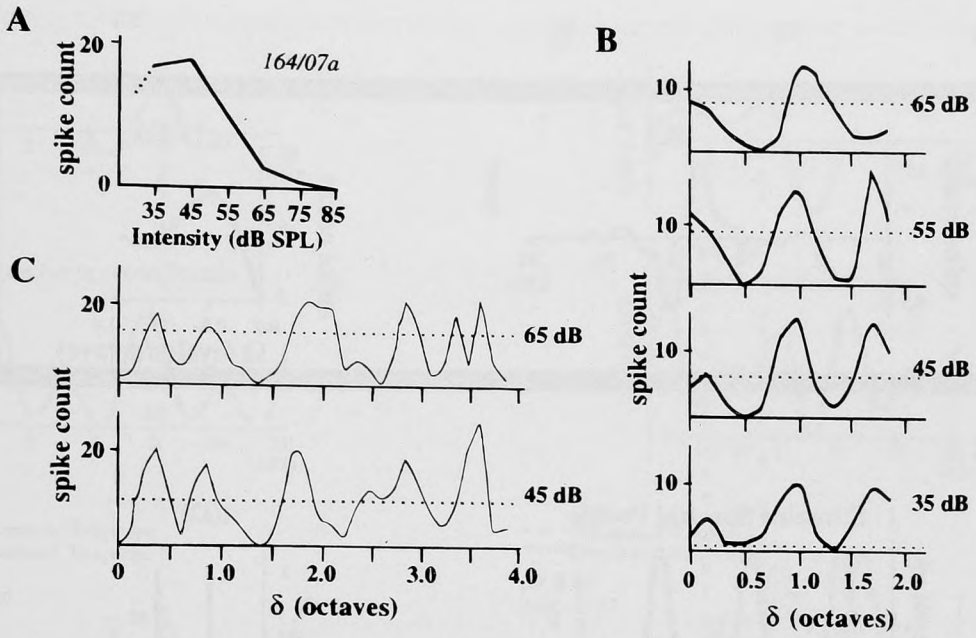



*FIGURE 4* **Responses for different overall stimulus levels.** (A) Nonmonotonic rate-level function of a single unit, measured between 35 and 85 dB SPL (dashed line represents extrapolated function below 35 dB SPL). (B) Measured response curves to a ripple pair spectral profile (0.4 and 0.8 cycle/octave) as a function of spectral shift. Absolute level of the stimulus is indicated to the right of each plot. (C) Measured response curves to a complex of five ripples for the same unit (0.4 to 2 cycle/octave in steps of 0.4 cycles/octave) as a function of profile shift. Absolute level of the stimulus is indicated to the right of each plot.

both cases, these changes are explained by the $T(\Omega)$ in that the predicted response $r_p(\delta)$ matches the basic features of the response curve $r_m(\delta)$.

### *Responses to Combinations of Three or More Ripples*

The responses of a unit to a progressively larger number of ripples is shown in Figure 7. In all cases, the match between predicted and measured responses is comparable ($\rho \approx 0.8$). The effect of the cell's filtering of different ripple amplitudes and phases is more dramatically seen with three or more ripples. For instance, in Figure 7B and C, the responses differ significantly from the shape of the spectral profile, which contains several ripples outside the $T(\Omega)$ band-pass.

Responses from two other cells to four and five ripple combinations are shown in Figure 8. Again, note the difference between the response curves and the corresponding spectral profiles. For instance, in Figure 8A, the relative strength of the responses to the stimulus peaks at 8 and 16 kHz is reversed; in Figure 8B, the response to the peak at 4 kHz (or approximately 30 kHz) is significantly narrower. In both cases, these response features are predicted from the RF (or the ripple transfer function). These examples, therefore, demonstrate that the responses to stimuli consisting of more than two ripple components basically superimpose as described for ripple pairs.

### *Summary of Responses to Ripple Combinations*

The results from all tests on AI units recorded are summarized in Figure 9. Figure 9A shows the distribution of the correlation coefficient between predicted and measured response, $\rho$, for ripple pairs. In 75% of all cells, fair predictions could be made ($\rho \geq 0.6$). Two of the worst three predictions belong to cells from the same penetration that had narrow transfer functions and relatively high characteristic ripples (1.6 cycles/octave). Figure 9B demonstrates that the correlation coefficient gradually decreases with the number of ripple components in the stimulus profile. Note that we have included in this plot for comparison correlation coefficients obtained with single ripples; these data are typical of the errors expected in measuring the transfer functions $T(\Omega)$.

### *Responses to Vowel-Like Spectral Envelopes*

Vowel spectra can be described in terms of ripple combinations of various amplitudes and phases as shown in Figure 10. Such complexes were presented consisting of 101 logarithmically spaced tones over 5 octaves (0.25 to 8 kHz), and with a spectral envelope constructed as a combination of 10 ripple components (0.2 to 2 cycles/octave). The responses were, as before, recorded as a function of shift of the spectrum ($\delta$) along the logarithmic frequency. Measured and predicted responses to the spectral profiles of the vowels /aa/ and /iy/ were

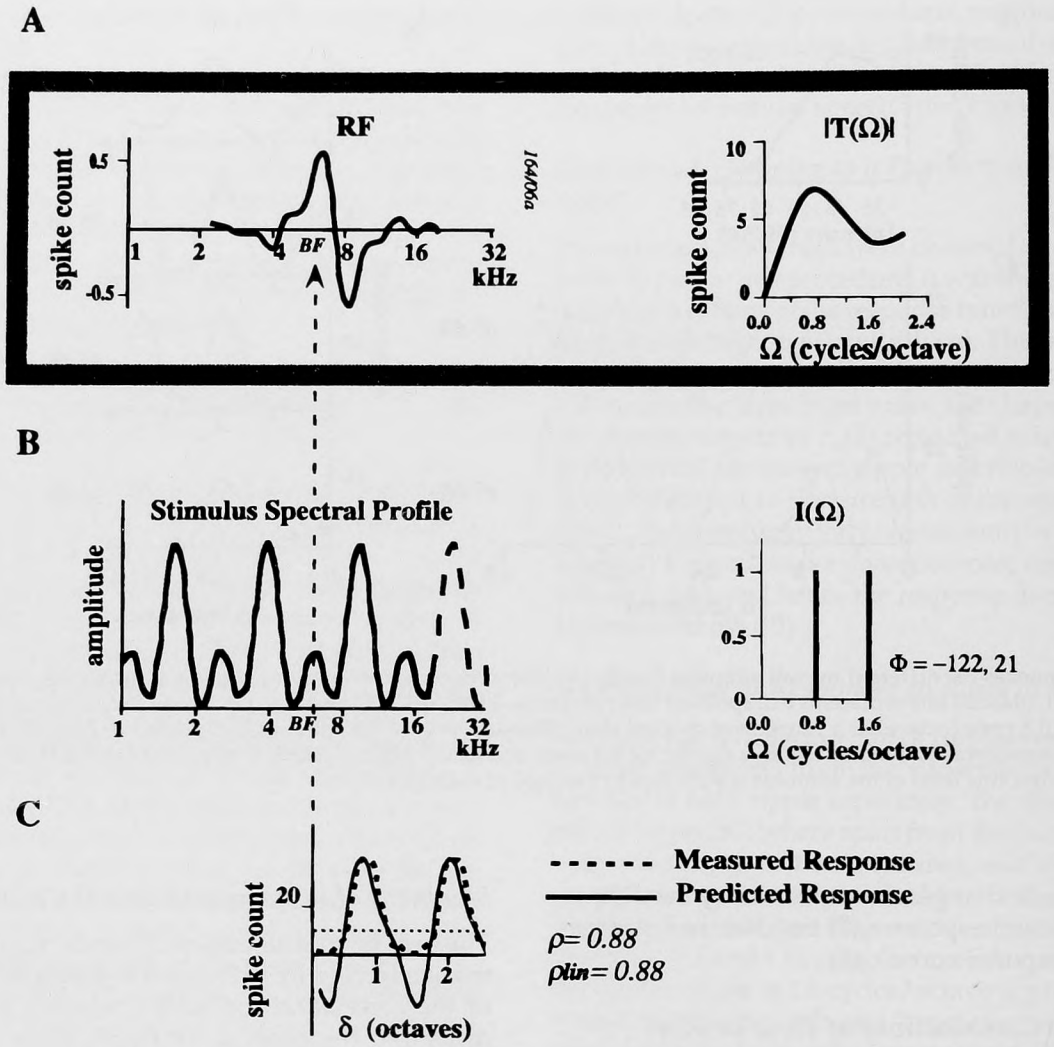



*FIGURE 5* **Responses of a single unit to a 2-ripple spectral profile.** (*A*) The RF and ripple transfer function of the unit. Such unit response properties are enclosed within a box in all remaining figures so as to highlight them. Details as described for Figure 2B,C. (*B*) The stimulus spectral profile and its ripple content, $I(\Omega)$. The dashed portion of the spectral profile is the (nonexistent) periodic extension of the profile, which is rotated in from the right as the profile is shifted by $\delta$. It is drawn simply to facilitate comparison with the measured and predicted response curves below. Details are as in Figure 3A. (C) Measured and predicted responses. Details are as in Figure 3C.

obtained in 8 units; two representative cases are shown in Figure 10A and B.

Measured and predicted responses are fairly matched in both cases. Furthermore, the responses differ significantly from the corresponding spectral profiles. For example, in Figure 10A, the unit responds vigorously only to the second peak of the /aa/ profile (at 1 kHz), presumably because of the one-sided inhibition seen in the RF of the cell. Similarly, the unit in Figure 10B responds better to the /iy/ peak at 3.5 kHz than to that at 0.3 kHz (or approximately 8 kHz) although the two are of equal height. These response features are quantitatively predicted from the transfer function of the units.

## DISCUSSION

We have examined here the extent to which AI cells respond linearly to their input spectral profiles. In another report (see Shamma *et al.*, 1995), it was concluded that a linear component must exist since parameters of the ripple transfer function were roughly correlated to those derived from the response area measured using tonal stimuli. In this study, a fundamental consequence of the linearity hypothesis is investigated, namely, the superposition principle. Specifically, it is shown that a unit response to a spectral profile composed of several ripples can be reasonably well predicted by the linear sum of its responses to the individual ripples, that is, from the

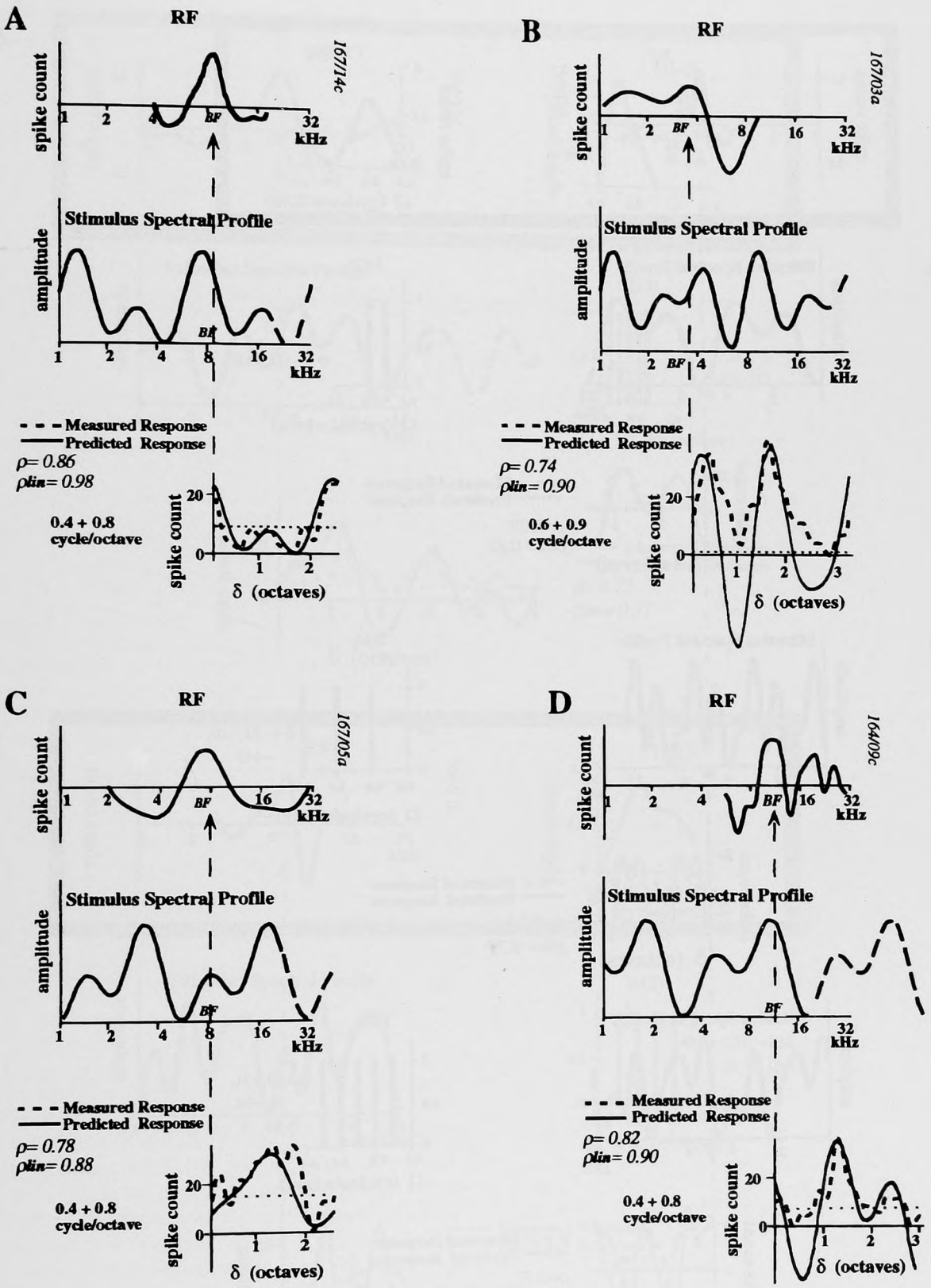



*FIGURE 6* **Measured and predicted responses to various ripple pairs for four different single units (A-D).** In each case, the RF of the unit is illustrated at the top, the stimulus spectral profile in the middle, and the responses at the bottom. The ripple content of the stimulus is indicated next to the bottom plot. Other details are as in Figure 5.

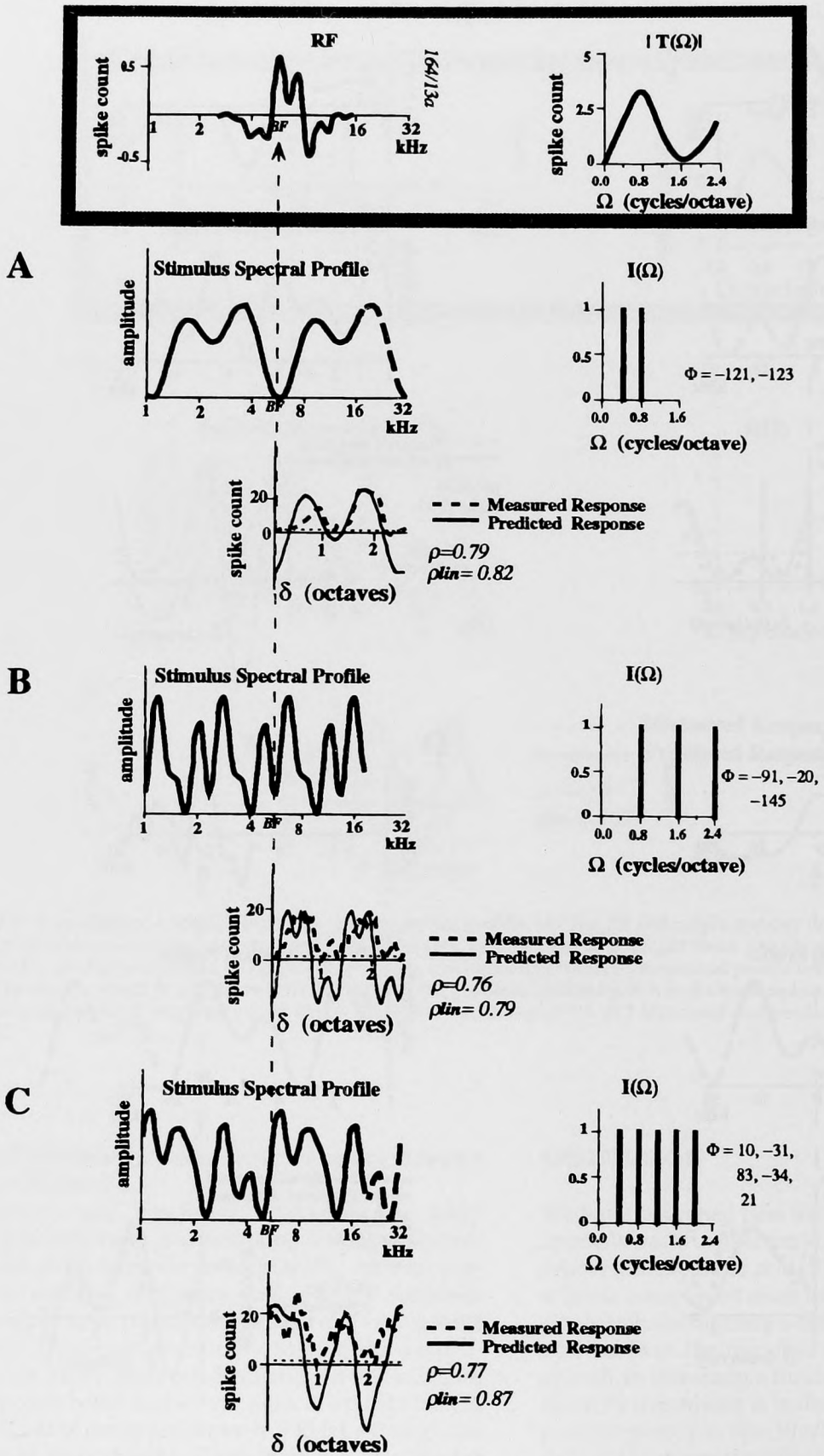



*FIGURE 7* **Measured and predicted responses of a single unit to stimulus profiles with increasing number of ripples (two in *A*, three in *B*, and five in C).** The response characteristics of the unit are shown in the box (top of figure). Other details are as in Figure 5.

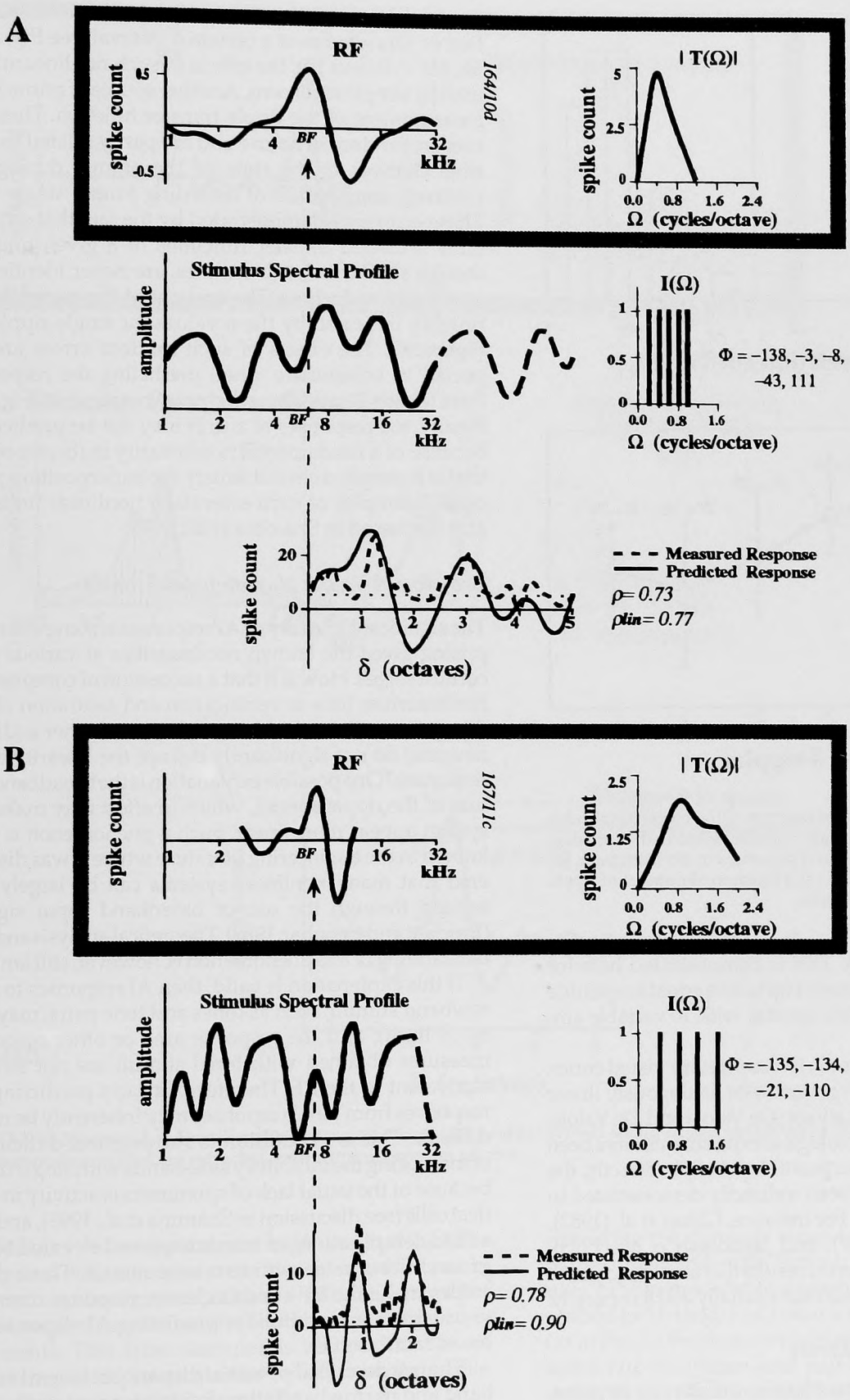



*FIGURE 8* **Measured and predicted responses of two units to multiple ripple stimuli.** Details are as in Figure 5. (A) Responses to profile of four ripple components. (B) Responses to profile of five ripple components.

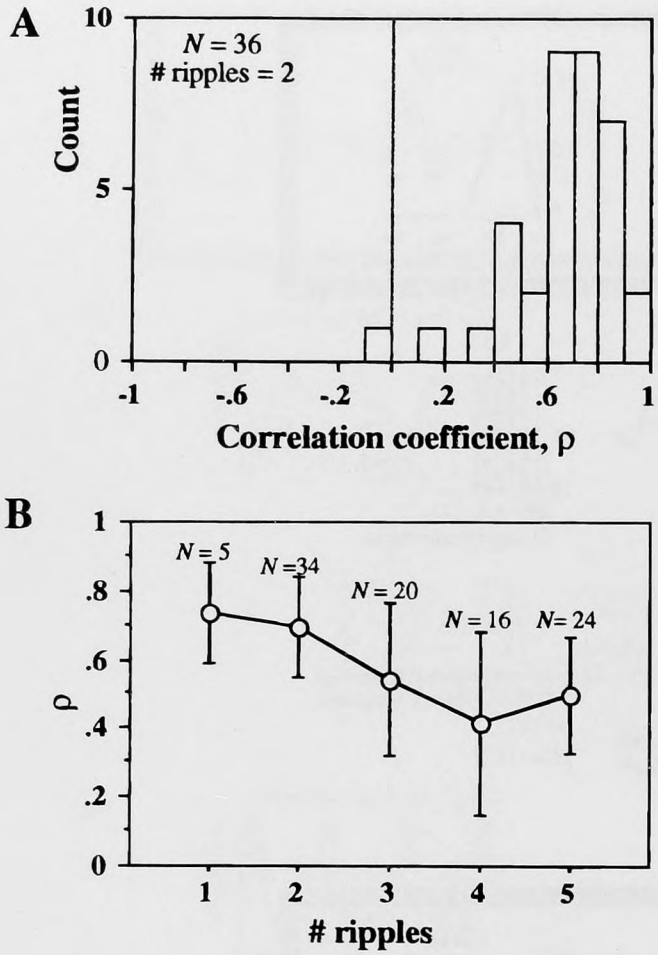



*FIGURE 9* **(A) Distribution histogram of the correlation index ρ of cells for two-ripple stimuli.** For cells tested with more than one ripple pair, the average ρ is represented in the histogram. (B) Mean and SD of ρ over all cells as a function of number of ripple components. N = number of cells.

ripple transfer function. This is demonstrated here for spectral profiles composed of up to five equal amplitude ripples, and for vowel-like spectra with 10 variable amplitude ripples.

Responses of simple cells in the primary visual cortex (VI) have also been interpreted to be analogously linear with respect to visual gratings (De Valois and De Valois, 1988). Although no physiological experiments have been reported to test the superposition principle directly, the linearity of VI cells has been indirectly demonstrated in a variety of other ways. For instance, Glezer *et al.* (1982), Jones and Palmer (1987), and Jagadeesh *et al.* (1993), among others, have obtained results that are strongly consistent with this hypothesis both spatially and temporally.

### *Sources of Prediction Errors*

Clearly, prediction errors (that is, differences between $r_p(\delta)$ and $r_m(\delta)$) can be found in all examples illustrated. They are attributable to various sources. For instance, measured responses are in many cases half-wave rectified or saturated over a certain δ interval (see Figs. 3, 5, 6B, 6D, 7, 8, and 10); the effects of such nonlinearities is usually simple to discern. Another source of errors is the measurement of the ripple transfer function. These errors are random in nature and are partly related to possible changes in the state of the animal during the relatively long period of recording from a single unit. These errors are demonstrated by the fact that sequentially recorded transfer functions of a given unit, although similar in basic outlines, are never identical in amplitude and phase. The amount of this variability is roughly indicated by the ρ values for single ripples in Figure 9B. The effects of such random errors are expected to accumulate when predicting the responses from increasing numbers of ripple components (Fig. 9B). Finally, the responses of a unit may not be predictable because of a fundamental nonlinearity in its responses, that is it simply does not satisfy the superposition principle. Examples of such essentially nonlinear units are also discussed in Shamma *et al.*, 1995.

### *Broadband Versus Narrowband Stimuli*

The significant linearity of AI responses is somewhat surprising given the known nonlinearities at various subcortical stages. How is it that a succession of compressive nonlinearities (due to rectification and saturation of auditory nerve fibers, cochlear nucleus, and other auditory neurons) do not significantly disrupt the linearity of AI responses? One possible explanation is the broadband nature of the ripple stimuli, which in effect may make the system appear more linear. Such a phenomenon is well known in the engineering literature where it was discovered that many nonlinear systems can be largely linearized through the use of broadband input signals (Brockett and Cebuhar, 1988). Theoretical analysis and understanding of this phenomenon is, however, still limited.

If this explanation is valid, then AI responses to narrowband stimuli, such as tones and tone pairs, may not be as linear, and the response area or other response measures obtained with tonal stimuli are not strictly equivalent to the RF. Therefore, linearly predicting AI responses from tonal responses may inherently be more difficult. This is in addition to the practical difficulties of measuring the inhibitory side-bands with single tones because of the usual lack of spontaneous activity in cortical cells (see discussion in Shamma *et al.*, 1995), and the added complications of interactions and elevated background firing rates with two-tone stimuli. These difficulties make the RF a much cleaner response measure to use than tonal stimuli for predicting AI responses to broadband profiles.

The just described potential disparity between broadband and narrowband stimuli may also explain the relative weakness of the correlations obtained between the RF and response area parameters (see Shamma *et al.*,

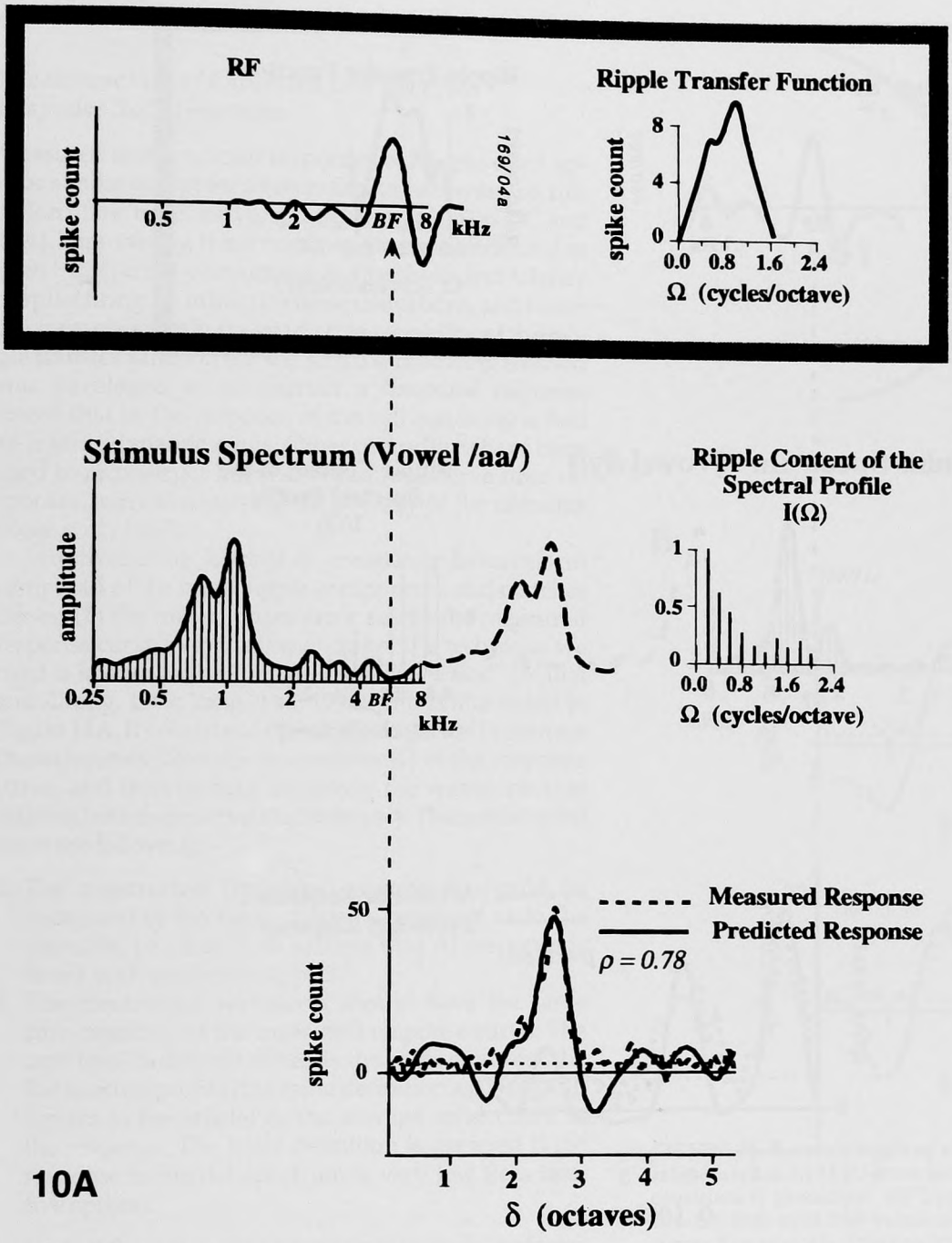
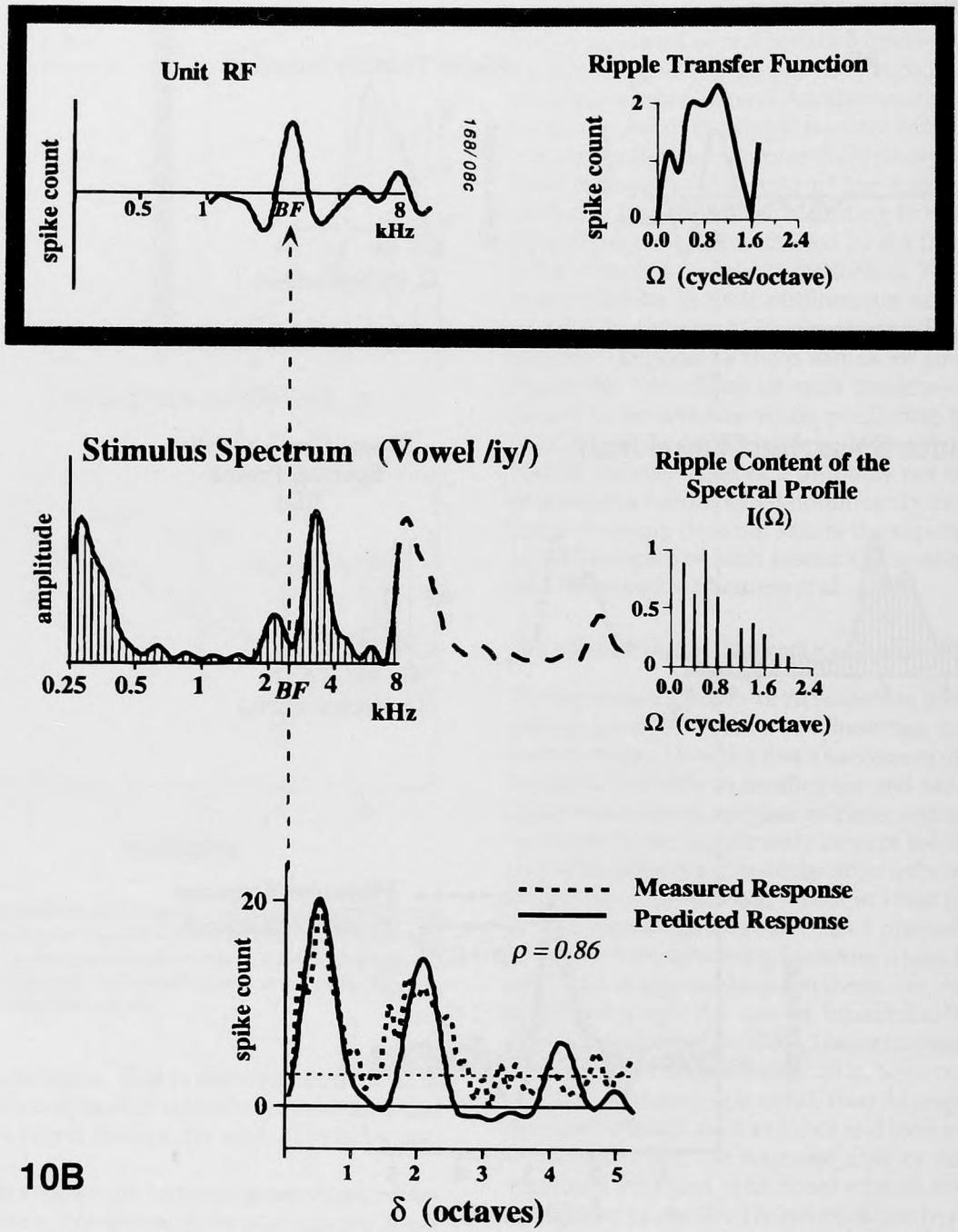



*FIGURE 10* **Measured and predicted responses to vowel-like profiles (vowel /aa/ in A, and vowel /iy/ in *B*).** All details are as in Figure 5. The stimulus profiles are extracted from naturally spoken tokens.

1995). A better correspondence might result if the RF is compared with response areas measured with tones in a broadband background, such as white noise or the flat tone complex used as carrier for the ripples in our experiments. This latter stimulus is identical to the so-called single increment profile widely used in profile analysis experiments (Green, 1988).

It should be emphasized here that the linearity of the responses observed in these experiments is not due to restricting the dynamic range of the input stimulus or of the output spike rate. Rather, it is seen for deep stimulus profile modulations (such as 90 to 100%, as described in Methods) and over a range of absolute levels (as in Fig. 4). Furthermore, simple nonlinearities such as spike rate saturation and half-wave rectification evidently do not affect the essential linearity of the response but rather limit our ability to "see" the full waveform, much like the way these nonlinearities affect the firing




rate of auditory nerve fibers (Rose *et al.*, 1967; see also the Appendix).

## SUMMARY OF MAJOR CONCLUSIONS

There are two conclusions implied by the results presented here and in Shamma *et al.*, 1995.

(1) Realistic sounds such as speech, music, and various environmental sounds are mostly broadband in nature. According to the experimental results, AI analyzes the acoustic spectrum of such sounds in a substantially linear manner.

(2) AI cells exhibit band-pass ripple transfer functions with a range of different characteristic ripples and phases. This suggests that AI does not represent the spectral profile directly, but, instead, it analyzes the profile into its constituent ripple components.

## APPENDIX

### *Reconstruction of Saturated and Rectified Response Rate Functions*

Measured and predicted responses of AI cells often appear similar except for a saturation or half-wave rectification of the measured response rate (as in Figs. 3C and 10A). Presumably, these nonlinearities are attributed to such biophysical phenomena as threshold and latency of spike firing. To minimize these distortions, and hence to assess objectively the predictive capability of the ripple transfer function (or the RF), the following method was developed to reconstruct a *linearized* response curve, that is, the response of the cell assuming it had an infinite dynamic range. Other procedures have been used to reconstruct linearized auditory nerve fiber responses, such as reversing the polarity of the stimulus (Rose *et al.*, 1967).

Intuitively, the algorithm constructs a waveform composed of the input ripple components and matches closest (in the mean square error sense) the measured response curve over the linear range. The technique we used is known as the "*convex projection method*" (Mallat and Zhong, 1989; Yang *et al.*, 1992), and is illustrated in Figure 11A. It consists of defining two sets of important characteristics (features or constraints) of the response curve, and then finding iteratively the waveform that satisfies both these sets simultaneously. The sets selected were the following:

1. The constructed (response) waveform should be composed of the same ripple components as in the stimulus, i.e., that is, to assume that AI response is linear with respect to ripples.
2. The constructed waveform should have the same zero crossings as the measured response curve. The zero level is defined either as the spike count for the flat spectral profile (the same definition as used in all figures in the article) or the average spike count of the response. The latter definition is prefered if the response to the flat spectrum is very low (less than five spikes).

Each of these two properties imply many waveforms (or *spaces* designated $S_1$ and $S_2$ in Fig. 11A). However, the conjunction of these two spaces of waveforms can be shown to define a unique waveform (Logan, 1977). To find it, we start with any arbitrary waveform ($w_1$) that satisfies one of these properties, that is, is formally in one space (such as, a square-wave that has the same level-crossings as the measured response). The waveform is then projected ($P_1$) onto the other space (find the closest curve ($w_2$) to the square wave that is composed only of the stimulus ripples). This latter curve ($w_2$) is now likely to have different zero crossings than the desired waveform. So now we repeat the procedure by projecting $w_2$ back unto $S_1$ ($P_2$), and so on until the projections yield a

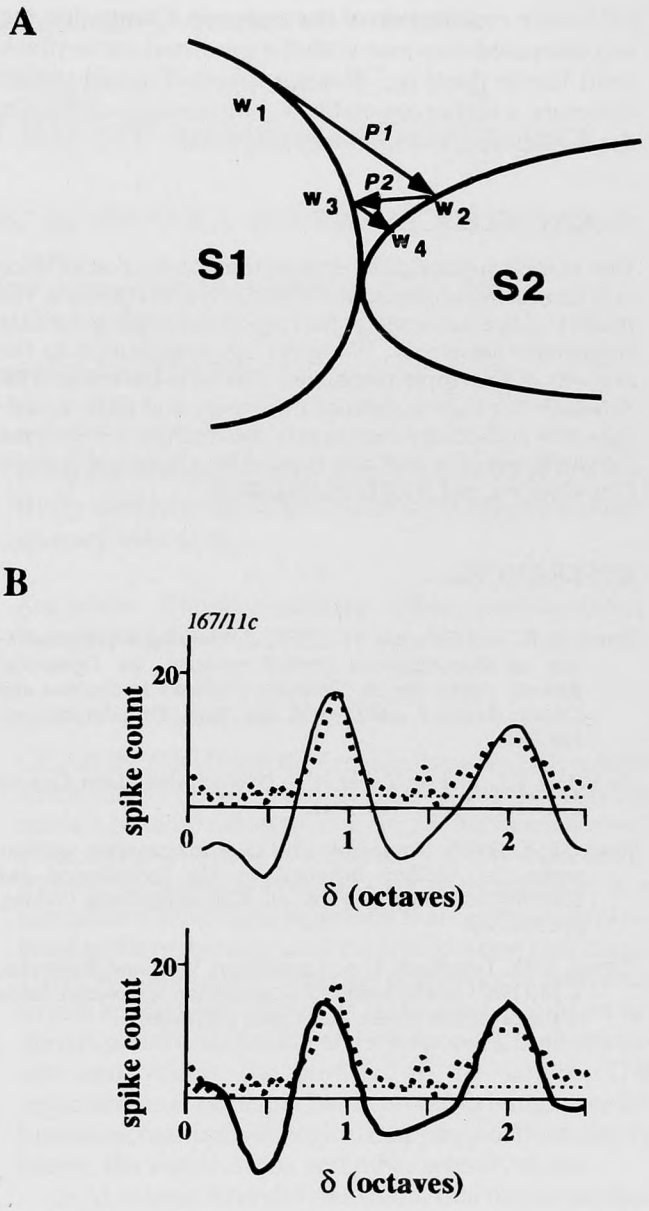



*FIGURE 11* **Reconstruction of measured responses to remove rate saturation and half-wave rectification.** (*A*) Schema of the reconstruction procedure. (*B*) Upper plot illustrates the reconstructed (thin solid line) versus measured (dashed line) response curves. Lower plot is of the predicted (thick solid line) versus measured response curves for the same test as in Fig. 8B. All details are as in Figure 3C.

stable (nonchanging) waveform. This procedure always converges for this problem because the two spaces selected are convex (Yang *et al.*, 1992). It usually takes no more than 20 iterations to find the desired waveform. A typical example of such a response reconstruction is shown in Figure 11B for the same unit and test as in Figure 8B. The reconstructed curve (thin solid line in upper plot) matches well the response curve above the baseline, and does not suffer from the

half-wave rectification of the response. Comparing the reconstructed response with the predicted curve (thick solid line in the lower plot; same as in Fig. 8B) yields, therefore, a higher correlation coefficient ($\rho_{lin} = 0.90$). On the average, $\rho_{lin}$ is larger than $\rho$ by 0.11.

## ACKNOWLEDGEMENT

This work is supported by grants from the Air Force Office of Scientific Research and the Office of Naval Research. We thank P. Gopalaswamy for his help in developing the data acquisition system, K. Wang for his contribution to the analysis of the ripple responses, and A. L. Owens and N. Kowalski for their assistance in surgery and data recordings. The authors are members of the Institute for Systems Research, which is partially funded by a National Science Foundation grant (NSFD CD 8803012).

## REFERENCES

Brockett, R., and Cebuhar, W. (1988). Smoothing and linearization of discontinuous control systems. In: *Dynamical Systems Approaches to Nonlinear Problems in Systems and Circuits*, Salam, F. and Levi, M. eds., Siam, Philadelphia, pp. 199–208.

De Valois, R.L., and De Valois, K. K. (1988) *Spatial Vision*. Oxford University Press, New York.

Evans, E. F. (1979). Single-unit studies of mammalian cochlear nerve. In: *Auditory Investigation: The Technological and Scientific Basis*, Beagley, H. A., ed., Clarendon Press, Oxford, pp. 324–367.

Glezer, V. D., Tsherbach, T. A., Gauselman, V. E., and Bondarko, V. M. (1982). Spatio-temporal organization of receptive fields of the cat striate cortex. *Biol. Cyber. 43*, 35–49.

Green, D. M. (1988). *Profile Analysis*. Oxford University Press, New York.

Jagadeesh, B., Wheat, H. S., and Ferster, D. (1993). Linearity of summation of synaptic potentials underlying direction selectivity in simple cells of the cat visual cortex. *Science 262*, 1901–1904.

Jones, J. P., and Palmer, L. A. (1987). The two-dimensional spatial structure of simple receptive fields in cat striate cortex. *J. Neurophysiol. 58*, 1187–1211.

Logan, B. (1977). Information in the zero-crossings of bandpass signals. *Bell Sys. Tech J. 56*, 510.

Mallat, S., and Zhong S. (1989). Complete signal representation with multiscale edges. Currant Institute, New York, Robotics Research Technical Report 483.

Phillips, D. P., Reale, R. A., and Brugge, J. F. (1991). Stimulus processing in auditory cortex. In: *Neurobiology of Hearing: the Central Auditory System*, Altschuler, R. A., Bobbin, R. P., Clopton, B. M., and Hoffman, D. W., eds., Raven Press, New York, pp. 335–365.

Plomp, R. (1976). *Aspects of Tone Sensation*. Academic Press, New York.

Rose, J. E., Brugge, J. F., Anderson, D. J., and Hind, J. E. (1967). Phase-locked responses to low-frequency tones in single auditory-nerve fibers of the squirrel monkey. *J. Neurophysiol. 30*, 760–793.

Schreiner, C. E., and Calhoun, B. M. (1995). Spectral envelope coding in cat primary auditory cortex. *Auditory Neurosci. 1*, 39–61.

Shamma, S. A., Fleshman, J. W., Wiser, P. R., and Versnel, H. (1993). Organization of response areas in ferret primary auditory cortex. *J. Neurophysiol. 69*, 367–383.

Yang, X., Wang, K., and Shamma, S. A. (1992). Auditory representations of acoustic signals. *IEEE Trans. Info. Theory 38*, 824–839.