Auditory Neuroscience, 1995, Vol 1, pp. 233–254
Reprints available directly from the publisher
Photocopying permitted by license only

Ripple Analysis in Ferret Primary Auditory Cortex. I. Response Characteristics of Single Units to Sinusoidally Rippled Spectra

SHIHAB A. SHAMMA, HUIB VERSNEL and NINA KOWALSKI

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

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

We compared the response properties of cells to tones and sinusoidally rippled spectral stimuli in the primary auditory cortex of the barbiturate-anesthetized ferret. Using broadband stimuli (1 to 20 kHz) with sinusoidally modulated spectral envelopes (ripples), the response magnitude of each cell was measured as a function of ripple frequency (Ω) and ripple phase (Φ), from which a "ripple transfer function" was constructed. Most cells (approximately 90%) responded best around a specific (characteristic) ripple frequency, Ω_o. Values of Ω_o range from 0.2 to 3 cycles/octave, with the average of the distribution around 1.0. Most cells also exhibited a linear ripple phase as a function of Ω. The intercept of the phase function is interpreted as the best (characteristic) ripple phase to drive the cell, Φ_o; the slope of the line reflects the location of the response area of the cell along the tonotopic axis. Φ_o ranges over the full cycle in a gaussian-like distribution around 0°. By inverse Fourier transforming the transfer function, we obtain the response field (RF) of the cell, an analogue of the response area measured with tonal stimuli. Parameters of the RF were compared to parameters of the tonal response area. The best frequency (BF) of the RF, BF_{RF}, was very similar to the tonal BF, and Ω_o and Φ_o were weakly but significantly correlated to the excitatory bandwidth and asymmetry index of the tonal response area, respectively. The RF was found to be a stable measure of a unit's response regardless of ripple amplitudes or overall stimulus levels. Responses to rippled spectra in auditory cortex closely resemble the response properties to sinusoidal gratings in the primary visual cortex. This provides a unified framework within which to interpret the functional organization of both cortices.

Key words: Primary auditory cortex, spectral ripples, ripple transfer function, response area, ferret

RESPONSE AREAS OF CELLS along the isofrequency planes of the mammalian primary auditory cortex (AI) have recently been shown to be systematically organized with respect to two properties: their excitatory bandwidths (Schreiner and Mendelson, 1990) and their asymmetry (Shamma et al., 1993). To measure the response areas, these investigations used simple tones that can be thought of as impulse-like stimuli along the tonotopic axis, as illustrated in Figure 1A. If cortical cells were to respond linearly, the measured response areas would reflect the "impulse responses" of the system along the tonotopic axis, and hence could be used to predict the system's responses to arbitrary spectra. Furthermore, by Fourier transforming the impulse response, one would obtain the corresponding transfer function, which represents the system's response to sinusoidally modulated spectra (Fig. 1B), more commonly known in the psychoacoustical literature as rippled spectra (Green, 1986). Consequently, response properties measured by tonal stimuli might be equally evident from their ripple transfer function.

The suggestion that cortical cells are linear might appear far-fetched given that their rate-level functions often exhibit threshold, saturation, and nonmonotonic behavior (Brugge and Merzenich, 1973; Pfingst and O'Connor, 1981; Nelken *et al.*, 1994). Nevertheless, just as measuring with tones a cell's bandwidth, tuning quality factor, or other linear systems response properties is considered meaningful, certain characteristics of the ripple responses may also prove useful, or possibly related to the properties measured with tones. It is possible as well that nonlinearities observed with tonal stimuli are less troublesome with broadband rippled

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

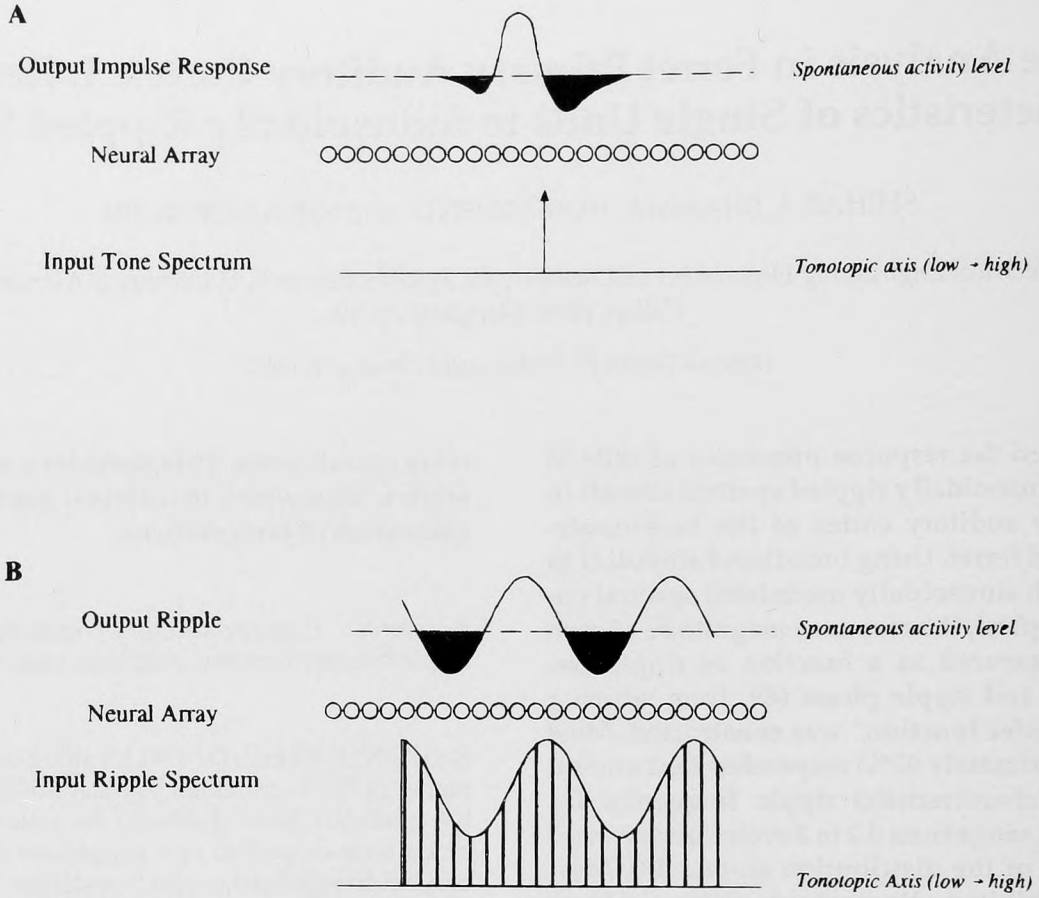


***FIGURE 1* Schematic of presumed responses of an array of AI cells uniformly distributed along the tonotopic axis.** (A) The responses to a single tone. The tone is represented by an impulse stimulus along the tonotopic axis. The tone evokes a pattern of cells' responses along the axis that mirrors the typical response area of a single cell with its excitatory tuned center and inhibitory sidebands (shaded region). The pattern is slightly asymmetrical, thus reflecting asymmetrical response areas of the array's cells. (B) The response pattern evoked by a rippled spectrum stimulus. The stimulus is represented by a sinusoid along the tonotopic axis. The output is an alternating (sinusoidal) pattern of excitation and inhibition, that is amplified or attenuated in amplitude, and phase-shifted relative to the input pattern. A transfer function can be measured by noting the amplitude and phase of the output relative to the input ripple at various ripple frequencies. As in *A*, the cells in this scheme have asymmetric response areas reflected here by a phase shift.

spectra, or negligible over a certain range of stimulus parameters (see Shamma and Versnel, elsewhere in this issue).

An analogous situation to the above has long existed in experimental studies of auditory nerve responses. There, nonlinearities such as firing rate rectification, saturation, two-tone suppression, and adaptation are prevalent (see review by Pickles, 1986). These nonlinearities, however, did not impede measurements of transfer characteristics of auditory-nerve fibers using single tones (Kiang *et al.*, 1965), noise stimuli (De Boer and De Jongh, 1978), or acoustic clicks (Pfeiffer and Kim, 1972), all implying strong linear components in the responses.

Our primary goal in this report is to measure the responses of AI cells to rippled spectra at various ripple frequencies and phases, that is, to measure their ripple transfer functions, and the dependence of this function on the amplitude of the ripples and the overall intensity of the sound. A second objective is to compare characteristic features of these transfer functions to response properties measurable using tonal stimuli, such as the bandwidth or the asymmetry of the response area.

Such an approach has proven fruitful in analogous studies of the primary visual cortex (De Valois and De Valois, 1988). There, transfer functions measured using sinusoidally modulated gratings reveal much about the functional organization of the system, and its response to more complex stimuli such as oriented bars. In auditory physiology, such stimuli have only been reported by Schreiner and Calhoun (1995). Recently, several psychoacoustical studies (Hillier, 1991; Summers and Leek, 1994; Vranić-Sowers and Shamma, 1995a,b) have converged on the similar notion that measuring the perceptual thresholds of rippled spectra may help explain how spectral profiles are perceived.

A somewhat different stimulus, called the "linear ripple" spectrum, has been used in psychoacoustical ex-

periments to investigate auditory-filter characteristics (Houtgast, 1977) and the perception of pitch (Yost *et al.*, 1978). It also has been used in physiological experiments in the dorsal cochlear nucleus (Bilsen *et al.*, 1975). Unlike our rippled spectra, which are sinusoidal along a logarithmic frequency axis (roughly the tonotopic axis), the linear ripple spectra are sinusoidal along the linear frequency axis, mimicking the harmonic structure of complex sounds. These spectra presumably do not evoke a sinusoidal excitation pattern along the tonotopic axis, which is roughly logarithmic (Greenwood, 1990); instead, they evoke a rippled pattern with an exponentially increasing ripple rate. Consequently, linear ripple spectra with appropriately weighted energy profiles (to compensate for the logarithmic crowding of higher frequency components) could theoretically be used to measure the transfer function of cortical cells on a linear frequency axis.

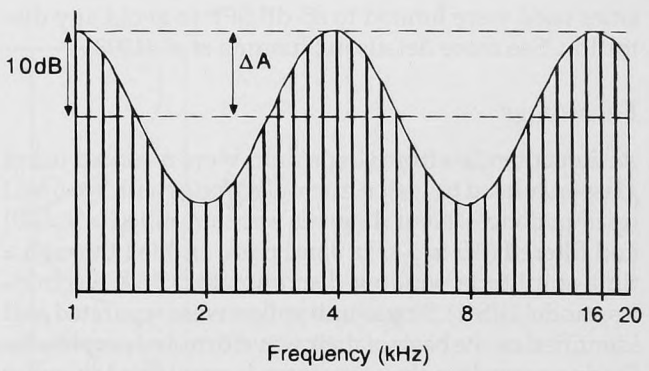


***FIGURE 2* Schematic of a rippled spectrum stimulus.** It is composed of 101 tones equally spaced along the logarithmic frequency axis between 1 to 20 kHz. The envelope is sinusoidally modulated on either a logarithmic (as in the figure) or a linear amplitude scale (not shown). The ripple phase is defined relative to a sine wave starting at the left edge. For the ripple shown, the ripple frequency $\Omega = 0.5$ cycles/octave, the ripple phase $\Phi = 90°$. The ripple amplitude (ΔA) is logarithmic at 10 dB.

METHODS

Surgery and Animal Preparation

The data presented here are obtained from experiments in seven young adult male ferrets, *Mustela putorius*, weighing approximately 1.5 kg. A detailed description of the animal preparation is given by Shamma *et al.* (1993). In brief, the ferrets were anesthetized with sodium pentobarbital (40 mg/kg). An areflexic level of anesthesia was maintained throughout the experiment by continuous intravenous infusion of pentobarbital (approximately 5 mg/kg/h) diluted with dextrose-electrolyte solution. The ectosylvian gyrus, which includes the primary auditory cortex (Kelly *et al.*, 1986) was exposed by craniotomy and the overlying dura was incised and reflected. The brain was covered in 2% agar in saline to reduce pulsations. The contralateral meatus was exposed, cleaned, and partly resected, and subsequently a cone-shaped speculum containing a Sony MDR-E464 miniature speaker was sutured to the meatal stump.

Acoustic Stimuli

Two types of stimuli were used in these experiments. The first type consisted of pure tone stimuli (single and two-tone bursts, 200 ms duration, 7 ms rise and fall times) and frequency modulation (FM) tones (2 octaves around the best frequency (BF), at sweeping rates 50 to 250 octaves/s, two sweep directions). These were generated using two independent function generators, gated and mixed, and then fed through a common equalizer into the earphone. Other parameters of the test stimuli are described in Results and in Shamma *et al.* (1993).

The second type of stimuli were broadband complex sounds. This stimulus, schematically shown in Figure 2, consisted of 101 tones that were equally spaced along the logarithmic frequency axis and spanning 4.32 octaves (such as 1 to 20 kHz or 0.25 to 5 kHz). The range was chosen so that the BF of the cell tested lay well within the stimulus' spectrum. The envelope of the complex was then modulated sinusoidally either on a logarithmic or a linear amplitude scale to create the so-called ripple spectrum. In the logarithmic scale case, the amplitude of the ripple was taken as the logarithm of the ratio of the amplitude of the peak of the spectral envelope to the baseline (left ordinate in Fig. 2). In the linear case, it was defined as the maximum percentage change in the component amplitudes. The overall level of the complex stimulus was defined by the level of a single frequency component, L_1 dB sound pressure level (SPL), in the flat complex. Thus, the overall level for a flat complex with 101 components (ripple amplitude ΔA at zero) was taken to be $L_1 + 10 \log(101) \approx L_1 + 20$ dB. The overall level was varied over a range of 30 dB. Higher levels were avoided to ensure the linearity of our acoustic delivery system.

The ripple frequency (Ω) is measured in units of cycles/octave against the logarithmic frequency axis (Fig. 2). The ripple phase (Φ) is measured in radians (or degrees) relative to a sine wave starting at the left edge (low-frequency edge) of the complex (Fig. 2). 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). Maximum tone inten-

sities used were limited to 85 dB SPL to avoid any distortion. See more details in Shamma *et al.* (1993).

Recordings

Action potentials from single units were recorded using glass-insulated tungsten microelectrodes with 5 to 6 MΩ tip impedances. Neural signals were amplified (×10,000) and filtered (0.5 to 5 kHz band pass) and led through a time-amplitude window discriminator (BAK electronics, model DIS-1). Single-unit spikes were separated and identified on the basis of their waveform and amplitude. During recording the waveform is monitored in order to check the stability of the spike waveform. 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 and spike count histograms of the responses.

In each animal, electrode penetrations were made orthogonal to the cortical surface. An experiment consisted of about 10 useful microelectrode penetrations. In each penetration, 1 to 7 cells were studied, typically at depths of 350 to 600 μm corresponding to cortical layers III and IV where excitatory phasic responses to single tones are strongest (Shamma *et al.*, 1993). At a given depth, several units were easily distinguishable based on their different spike shapes and amplitudes in the recording trace. When an electrode is moved over 100 μm, isolated units were assumed to be different unless physiological response parameters were found to be unusually similar.

Data Analysis for the Tonal Stimuli

For each cell, we first manually determined the BF, defined as the frequency of the lowest threshold, followed by an isointensity response curve with up to 1/8 octave resolution at a low stimulus level. The rate-level function at BF was measured to determine the cell's response threshold and the nonmonotonicity, that is, whether the spike count decreased by more than 25% as intensity was increased.

Subsequently, the response area was determined using the two-tone stimulus as described in detail in Shamma *et al.* (1993). In brief, it consisted of two tone bursts of equal duration. The first tone burst (T1) was presented at several different frequencies centered around the BF of the cell to measure the excitatory response area. Since many cells exhibit low spontaneous firing rates, a second tone burst (T2) was fixed at BF (about 20 dB above threshold), with a 50 ms intertone delay to provide a level of background activity against which the inhibitory response area could be measured. Staggered onset times for T1 and T2 were used to segregate the phasic responses to the two tones, which made it relatively easy to determine in the same test the borders of the excitatory responses to T1 and of the inhibitory influences on T2. Note that the term "response area" is used here to denote the response of a cell as a function of a tone's frequency and intensity. In the experiments reported here, responses were usually obtained only at two T1 intensities (equal to or 10 dB higher than T2); thus in using the term "response area" reference is made only to specific slices of the area.

The bandwidth of the excitatory response area was determined at 20 dB above the threshold at BF, and denoted as BW20. The upper and lower frequencies were found by interpolation using a criterion of 10% of the maximum response. A correction was made for possible spontaneous activity.

Another important feature of the response area is the asymmetry of its inhibitory and excitatory portions around the BF. To quantify this feature, the following simple statistic was introduced in Shamma *et al.* (1993):

$$M = \frac{R_{>BF} - R_{<BF}}{R_{>BF} + R_{<BF}} \tag{1}$$

where $R_{>BF}$ and $R_{<BF}$ are the total number of spikes to both tones for an equal number of frequencies above and below the BF, respectively. If the excitatory and inhibitory responses are approximately symmetrical around the BF, the measure (M) will be near zero. Inhibition of T2 responses by T1 stimuli above BF or spread of T1 excitation to lower frequencies (below BF) causes M to be negative. Conversely, stronger low-frequency inhibition or high-frequency T1 excitation produces positive M values. Note that the M index as first defined is computed at one T1 intensity. Since the two-tone test was performed at two intensities of T1, indices are computed for both intensities and then averaged.

Finally, FM tones were presented in two sweep directions, and at different rates and at similar intensities as in the two-tone tests. The symmetry of the average responses to the two sweep directions was assessed using the index C as follows (Mendelson and Cynader, 1985; Shamma *et al.*, 1993):

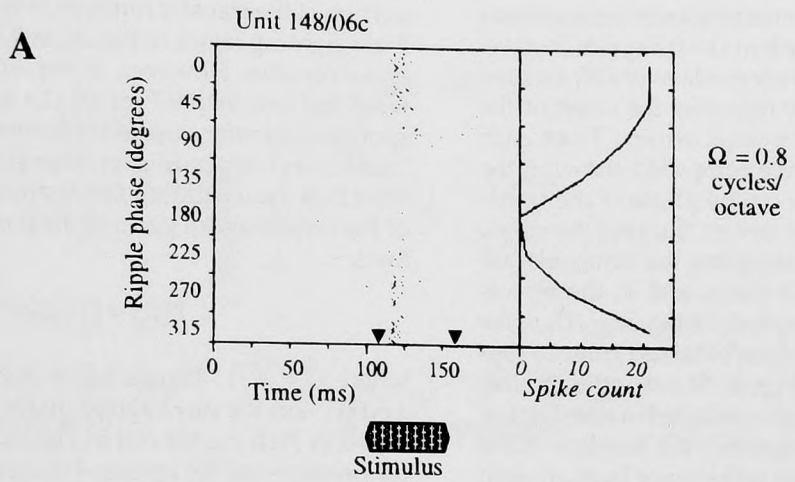
$$C = \frac{R\downarrow - R\uparrow}{R\downarrow + R\uparrow} \tag{2}$$

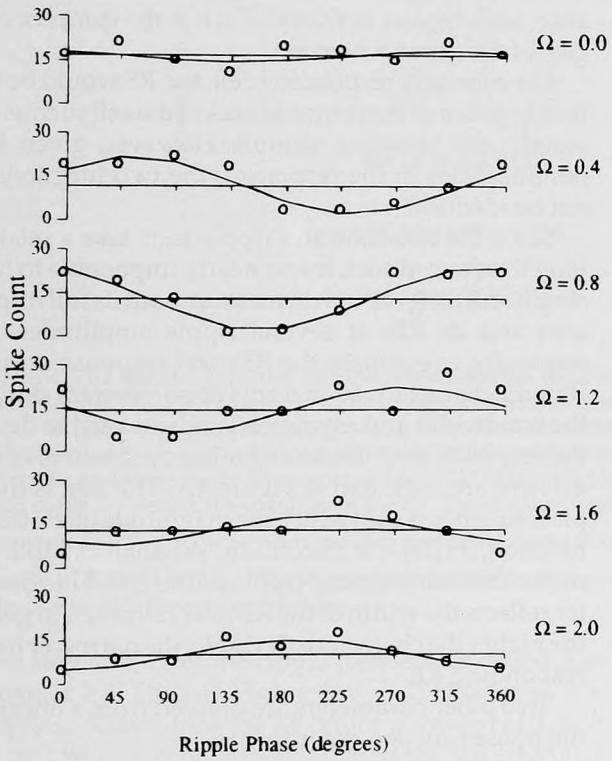
where $R\uparrow$ and $R\downarrow$ are the spike counts to the up and down sweeps, respectively. The index is computed for each intensity presented and then averaged.

Data Analysis for the Rippled Spectrum Stimuli

Following the tonal stimuli, a series of tests were carried out using rippled spectra with a range of ripple frequencies Ω (typically, from 0 to 4 cycles/octave with different resolutions) and ripple phases Φ (from 0 to $7\pi/4$ in $\pi/4$ steps). Different ripple amplitudes and overall stimulus levels were also tested.

Figure 3 illustrates the display and initial analysis applied to the data. Here the cell was tested over ripple frequencies 0 to 2 cycles/octave in steps of 0.4 cy-





C

Ω	DC(Ω)	AC_1(Ω)	AC_2(Ω)	AC_3(Ω)	AC_4(Ω)
0.0	16.25	1.98 129°	2.47 45°	3.29 -112°	1.00 -90°
0.4	11.62	9.81 30° (- 2π)	2.30 -139°	2.41 88°	0.25 -90°
0.8	12.62	11.06 111° (- 4π)	3.13 -61°	0.57 -32°	2.25 90°
1.2	14.38	7.78 174° (- 6π)	5.15 129°	4.17 141°	3.75 -90°
1.6	12.75	5.15 -124° (- 6π)	3.95 -55°	0.45 139°	3.00 -90°
2.0	9.88	5.24 -98° (- 8π)	1.03 -14°	1.40 58°	3.25 -90°

FIGURE 3 **The analysis of the responses to rippled stimuli.** (A) Raster of the responses of an AI cell (148/06c; BF of 7.5 kHz) to a rippled spectrum stimulus (Ω = 0.8 cycles/octave) at various ripple phases (0° – 315° in steps of 45°). The stimulus burst starts at 100 ms and lasts for 50 ms. Stimulus is repeated 20 times for each ripple phase. Spike counts as a function of the ripple are computed over a 50 ms window as indicated by the bold arrows and are displayed in the inset plot to the right of the raster. (B) Spike counts as a function of ripple phase for various ripple frequencies Ω between 0 and 2 cycles/octave. At each Ω, spike counts are indicated by the circles, and the abscissa is placed at the spike count averaged over all phases. The solid line is the best sinusoidal fit to the points (in the sense of mean square error). (C) An eight-point Fourier transform of the spike counts at each Ω yields estimates of the average spike count, $DC(\Omega)$, and the amplitude and phase of the best sinusoidal fit, $AC_1(\Omega)$, and of the 2nd, 3rd, and 4th (distortion) harmonics of the fundamental (Ω). Phase values computed by the fast Fourier transform algorithm are usually restricted between π and –π (wrapped). The phases of the AC_1 components can be unwrapped by adding the integer multiple of 2π indicated in the parentheses.

cles/octave. In Figure 3A, the raster shows the responses to different phases of the ripple at $\Omega = 0.8$ cycles/octave. The response spike counts were made over a 50 ms time window starting shortly (10 ms) after the onset of the stimulus, as indicated by the short arrows. From each such raster, spike counts were computed showing the dependence of the response on the phase of the ripple (inset plot to the right of the raster). Varying the ripple phase here is analogous to sampling the sinusoidal response of a filter at different times, and so these plots play the same role as the period histogram. All spike counts reflect the total responses obtained from 20 repetitions of each stimulus. In Figure 3B, such spike counts from all ripple frequencies are combined in one display. Note that for each ripple frequency the baseline of the plot is set equal to the average spike count from all eight phases.

To analyze the different components of the response, and specifically to estimate the level of the component synchronized to each ripple frequency, an eight point Fourier transform as performed on the spike counts at each ripple frequency (Fig. 3C). The first column is the baseline or the DC component in the responses to each ripple (or simply the average spike count). The second column (labeled $AC_1(\Omega)$) lists the magnitude and phase of the primary response component synchronized to the ripple frequency Ω. In effect, this is the amplitude and phase of the sinusoid that best fits the data at each ripple frequency (solid lines in Fig. 3E). Higher-order harmonics of the response to each ripple are listed in the other columns.

If the cell responses were completely linear and noise-free, $AC_1(0)$ (the synchronized component to a flat spectrum) and high-order harmonics at all Ω would vanish. In that case, $AC_1(\Omega)$ is the transfer function of the cell. However, this is rarely the case. So in order to take into account the magnitude of the nonlinearity and to reduce the effects of statistical noise, a weighted measure of the $AC_1(\Omega)$ responses [called $T(\Omega)$] is defined as follows:

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

$$T(\Omega) = 0 \text{ if } |AC_1(\Omega)| - |AC_1(0)| < 0 \tag{3}$$

where $|AC_1(\Omega)|$ is the magnitude of the response $AC_1(\Omega)$. The term $AC_1(0)$ reflects an estimate of noise, and the weighting factor $\left(\frac{|AC_1(\Omega)|}{\sqrt{\Sigma_{i=1}^{4} |AC_i(\Omega)|^2}}\right)$ reflects the quality of the sinusoidal fit and hence the linearity of the response. Weighting values larger than 0.7 indicate a relatively strong linear component, that is, $AC_1 > \sqrt{\Sigma_{i=2}^{4} |AC_i|^2}$. This is the case for all responses at $\Omega > 0$ in Fig. 3C. Typical values near the central (tuned) portion of the transfer function exceed 0.8. For example, the weighting factor in Fig. 3C is 0.94 at $\Omega = 0.4$ and 0.8 cycles/octave. However, if the $AC_1(\Omega)$ component is weighted less than 0.7 for all Ω's tested then the corresponding transfer function is termed "unresponsive" or "nonlinear," depending on other factors discussed later. Note that the weighting factor does not affect the phase of the responses. In general, $T(\Omega)$ can be written as follows:

$$T(\Omega) = |T(\Omega)| e^{j\phi(\Omega)} \tag{4}$$

where $j = \sqrt{-1}$. Figure 4A illustrates the magnitude $|T(\Omega)|$ and the unwrapped phase $\Phi(\Omega)$ of the transfer function $T(\Omega)$ for the cell in Figure 3. This ripple transfer function can be inverse Fourier transformed to obtain the impulse response. The reverse of this function will be called the response field (RF) of the cell shown in Fig. 4B and designated below as $w(x)$:

$$w(x) = \mathfrak{F}^{-1}\{T^*(\Omega)\} \tag{5}$$

where $F\{\cdot\}$ designates the inverse Fourier transform operation with respect to Ω, and $T^*(\Omega)$ is the complex conjugate of the transfer function.

For a linearly responding cell, the RF would be identical to a slice of the response area of the cell such as measured with two-tone stimuli. However, given likely nonlinearities in the responses, the two functions will not be identical.

Since the two-tone and ripple tests take a relatively long time to conduct, it was nearly impossible to hold a single unit long enough to measure both its full response area and its RFs at several ripple amplitudes. Consequently, to compare the RFs and response areas, the RF was characterized in terms of parameters similar to the bandwidth and asymmetry indices used to describe the response area. These and other parameters of interest here are indicated in Figure 4A. The first is the ripple frequency (Ω_0) at which the magnitude of the transfer function, $|T(\Omega)|$, is maximum. We shall call this parameter the characteristic ripple of the cell. 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

$$\Phi(\Omega) = x_0\Omega + \Phi_0 \tag{6}$$

where x_0 is the slope of the line ($= \frac{\Phi}{2\pi\Omega}$ in octaves), and Φ_0 is its intercept. The parameter x_0 reflects the location of the center of the RF relative to the left edge of the ripple (Fig. 4B). The term $k \cdot \frac{2\pi}{\Delta}$ resolves an unavoidable ambiguity in this estimate due

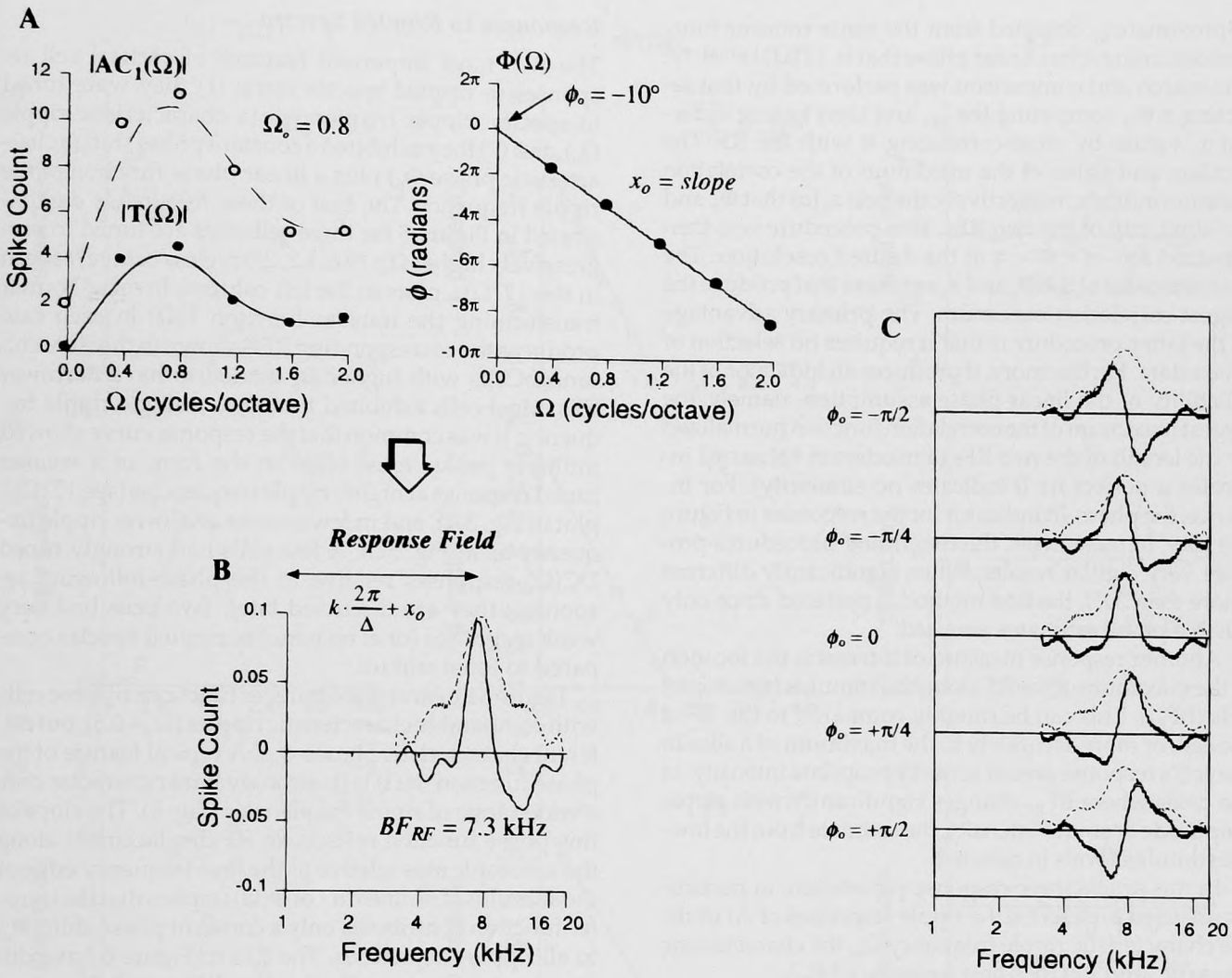


***FIGURE 4* The transfer function and the response field.** (A) The transfer function $T(\Omega)$ derived from the responses in Figure 3. Plot to the left is of the $|AC_1(\Omega)|$, the magnitude of the fundamental component of the Fourier analysis in Figure 3C. $|T(\Omega)|$ is a weighted version of $|AC_1(\Omega)|$. The characteristic ripple Ω_0 is the location of the maximum of $|T(\Omega)|$. Plot to the right depicts the phase function $\Phi(\Omega)$ (filled circles) of the fundamental component $AC_1(\Omega)$ (or of the transfer function $T(\Omega)$). The solid line represents a linear fit to the data with intercept Φ_0 and slope x_0. (B) The response field (RF) of the cell derived by an inverse Fourier transform of the transfer function $T(\Omega)$. The dashed curve represents the envelope of the RF. The distance from the left edge of the rippled spectrum to the center of the RF envelope is given by $k \cdot \frac{2\pi}{\Delta} + x_0$, where Δ is the step size of the ripple frequencies Ω tested, and k is an integer ≥ 1. The location of the maximum of the RF is defined as BF_{RF}. (C) Several RFs derived from the same $|T(\Omega)|$ and x_0 as in *A*. but with different Φ_0s. Changing the Φ_0s results in RFs of different asymmetries.

to the step size, Δ, of the ripple frequencies tested; k is an integer ≥ 1. Thus, in Figure 4, $x_0 = -2.14$ octaves, $\Delta = 0.4$ cycles/octave, then the RF center is located at 0.36 ($k = 1$) or 2.86 ($k = 2$) octaves relative to the left edge at 1 kHz (approximately 1.3 or 7.3 kHz). Using smaller Δ (such as 0.2 cycles/octave) or simply determining roughly the BF of the unit disambiguates the choice. The center of the RF is defined as the center of the envelope of the RF (dashed line in Fig. 4B); this envelope is computed from the analytic signal corresponding to the RF function (Oppenheim and Schafer, 1990).

The parameter Φ_0 shall be called the characteristic phase. It roughly reflects the asymmetry of the RF about its center. For instance, the RF is symmetrical for $\Phi_0 = 0$, and strongly asymmetric for $\Phi_0 = \pm 90°$. This is illustrated in Figure 4C, where the RF of the cell is computed for different Φ_0s by simply sliding its fitted phase function $\Phi(\Omega)$ vertically.

The straight line fit of the phase data was done in two ways. The first was to select the reliable phase values (such as those for which the weighting factor $\left(\frac{|AC_1(\Omega)|}{\sqrt{\Sigma_{i=1}^{4} |AC_i(\Omega)|^2}}\right)$ is greater than 0.7), and then to fit these points using regular least square error procedures. The second method was to compare the RF obtained from the transfer function, $|T(\Omega)|\, e^{j\Phi(\Omega)}$, to the

approximate $_{RF}$ obtained from the same transfer function assuming it has linear phase that is $| T(\Omega) | e^{j(x_o\Omega+\Phi_o)}$. The search and comparison was performed by first selecting a Φ_o, computing the $_{RF}$, and then testing different x_o values by cross-correlating it with the RF. The location and value of the maximum of the correlation function indicate, respectively, the best x_o for that Φ_o, and the similarity of the two RFs. This procedure was then repeated for $-\pi < \Phi < \pi$ at the desired resolution. The best estimates of the Φ_0 and x_0 are those that produce the largest correlation maximum. The primary advantage of the latter procedure is that it requires no selection of phase data. Furthermore, it produces an indicator of the reliability of the linear-phase assumption, namely, the largest maximum of the correlation function normalized by the length of the two RFs (a maximum value of 1 indicates a perfect fit; 0 indicates no similarity). For instance, the phase-fit indicator for the responses in Figure 4 is 0.99. In most cases, the two fitting procedures produce very similar results. When significantly different (more than 30°), the first method is prefered since only reliable phase estimates are used.

Another response measure of interest is the location of the maximum of the RF along the stimulus (tonotopic) axis (BF_{RF}). This can be roughly compared to the BF of the cell, or more accurately to the maximum of a slice in the cell's response area at some appropriate intensity. In the cases where BF_{RF} changes significantly with ripple amplitude or sound intensity, the estimate from the lowest stimulus levels in selected.

In this article three response parameters in particular are used to describe the ripple responses of AI units: the characteristic ripple frequency Ω_o, the characteristic phase Φ_o, and the RF's best frequency BF_{RF}.

RESULTS

The data illustrated here were collected from a total of 104 single-unit recordings in seven animals. All these units responded to tones. For 92 of tested AI units (88%) responses to rippled spectra were locked to the phase of the stimulus. An example of dependence on phase is shown in Figure 3A, where for a 180° shift the response changes from maximum to minimum (zero). In this article the responses are considered with respect to their phase-following (or vector-strength) of the ripple. The response irrespective of phase (for example, the average rate) is only discussed in the context of nonlinear behavior.

In this section detailed features of the responses to the rippled spectra are first described, together with their dependence on the ripple amplitude and overall sound level. Next, the correspondence between the shapes of the RFs and of the response areas is examined in general terms, and then in specific terms using such parameters as the characteristic ripple and phase, the response area bandwidth, and asymmetry indices.

Responses to Rippled Spectra

The two most important features of cortical cell responses to rippled spectra were: (1) they were tuned to specific ripple frequencies (a characteristic ripple Ω_o), and (2) they exhibited a constant phase shift (a characteristic phase Φ_o) plus a linear phase function of the ripple frequency. The first of these features is demonstrated in Figure 5 for three cells that are tuned to progressively higher Ω_os (0.6, 1.2, 2.0 cycles/octave) as seen in the $| T(\Omega) |$ plots in the left column. Inverse Fourier transforming the transfer function $T(\Omega)$ in each case produces the corresponding RFs (shown in the right column). Cells with higher Ω_0 tended to have narrower RFs. Most cells exhibited tuning to a single ripple frequency. It was common that the response curve showed multiple peaks, most often in the form of a weaker tuned response at higher ripple frequencies (see $| T(\Omega) |$ plot in Fig. 3A), and in fewer cases at a lower ripple frequency as in Fig. 5C). A few cells had strongly tuned DC(Ω) responses relative to the phase-following responses; they are discussed later. Two cells had very weak responses (or even none) to rippled spectra compared to tonal stimuli.

Figure 6 illustrates the transfer functions of three cells with comparable characteristic ripples ($\Omega_0 \approx 0.8$), but different characteristic phases Φ_0s. A typical feature of the phase function $\Phi(\Omega)$ is its strongly linear character over a wide range of ripple frequencies (Fig. 5). The slope of this phase function reflects an RF displacement along the tonotopic axis relative to the low-frequency edge of the stimulus. The linearity of $\Phi(\Omega)$ implies that the transfer function contributes only a constant phase shift (Φ_0) to all ripple frequencies. The RFs in Figure 6 have different asymmetries reflecting the different Φ_0's.

To assess the accuracy of fitting the phase function by a straight line, we computed the normalized match between the RF of the cell (Fig. 4B), and the reconstructed RF computed from the same transfer function, assuming the phase function is linear (Fig. 4C). If the two patterns are identical, the indicator is 1 and if they are random the indicator is 0. As is evident in all examples in Figures 4 through 6, the linear fit is accurate. In 84% of all cells responding to the ripples, the indicator exceeded 0.9. That is, in these cases, the assumption of a linear phase is justified, and the Φ_0 and x_0 are both meaningful measures of the RF.

The distributions of characteristic ripples and phases in our sample are shown in Figure 7A and B. The range of Ω_0 is 0.2 to 3 cycles/octave, and the average of the distribution lies around 1.0. The distribution of Φ_0 is gaussian-like with most cells (61%) between –45 and +45. The distribution is slightly biased toward negative values. In Figure 7C-D, scatter plots of the Ω_0 and Φ_0 against the BF of each unit are shown. The only apparent dependence on BF is seen for the Ω_0's where units with a lower BF tend to have smaller values. The joint distribution of these

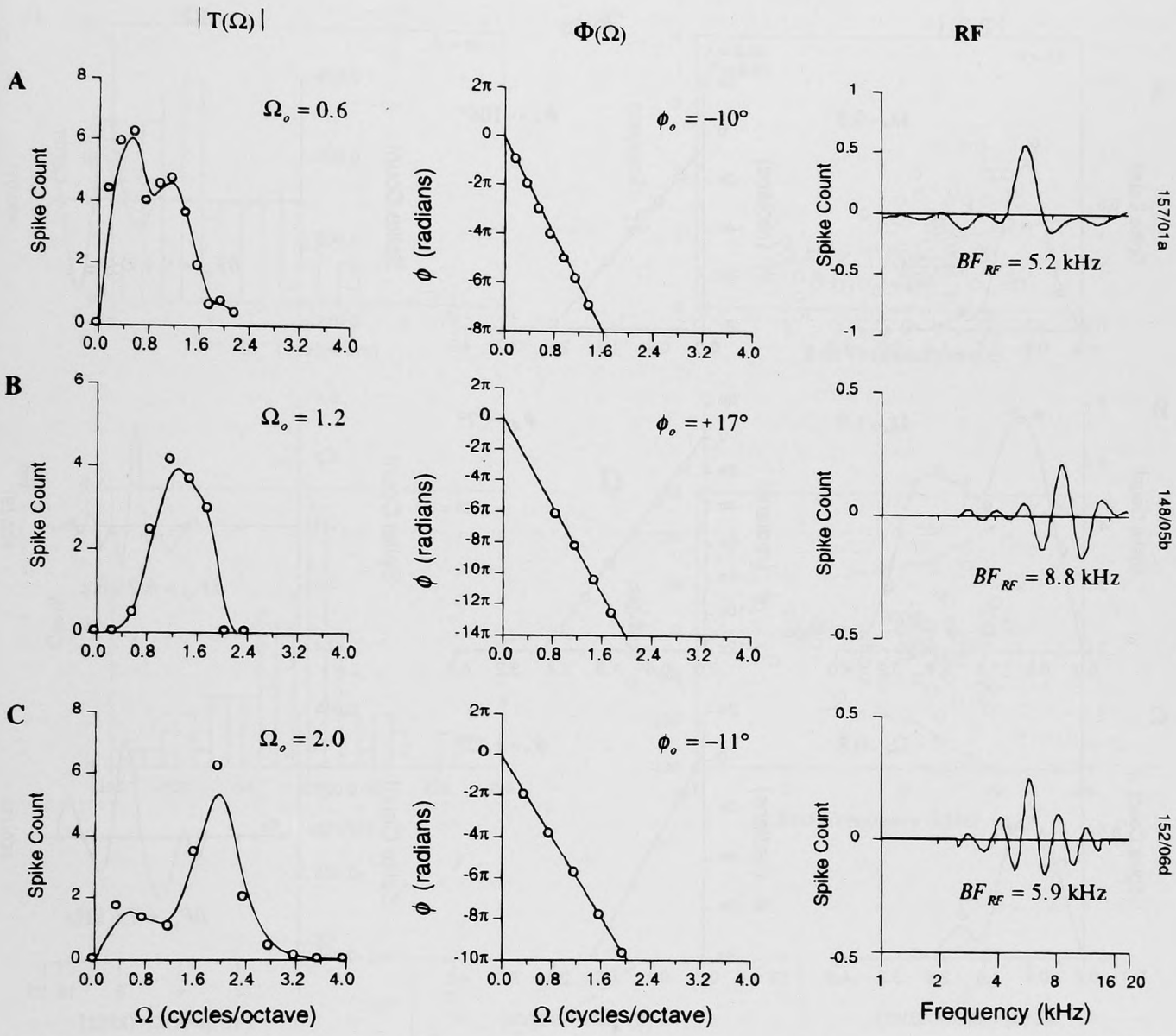


FIGURE 5 **Examples of ripple responses from three cells (*A, B, C*) with different characteristic ripples (Ω_o = 0.6, 1.2, 2.0 cycles/octave, respectively), but similar asymmetries (Φ_o = −10°, 17°, −11°).** For each cell, the left plot depicts the magnitude of its ripple transfer function $|T(\Omega)|$. Data points are connected by straight line segments and are then smoothed by linearly interpolating the slopes between the midpoints of adjacent segments. The middle plot represents the phase function $\Phi(\Omega)$, and the linear fit to the data points. The phase-fit indicators for the three cells are 0.99, 1.00, and 0.99. The right plot illustrates the corresponding RF.

two response measures is reflected by the scatter plot in Figure 7E. It indicates that the mean and variance of the Φ_0 distribution remain relatively constant and hence independent of the Ω_0. Finally, the average width of the transfer functions $|T(\Omega)|$ against the Ω-axis (measured at the 3 dB points) is found to be 1.7 octaves (SD = 0.9).

Linear Versus Logarithmic Ripple Spectra

It is unknown whether the central auditory system encodes the shape of the acoustic spectrum on a linear or a logarithmic amplitude scale, or via other representations such as the power spectrum. Because of this uncertainty, the responses of nine cells were measured using both linear and logarithmic ripples in order to determine whether any of the response features depended critically on this choice. There are no systematic differences between the responses to these stimuli, as verified in a paired t test on the parameters BF_{RF}, Ω_o, Φ_o, and the width of $|T(\Omega)|$ ($t < 1.1$; n = 9; p >0.1). The similarity between responses to ripples of either amplitude mode is illustrated for two units in Figure 8. Slight differences in magnitude or phase functions might show up without essentially affecting the estimate of the RF.

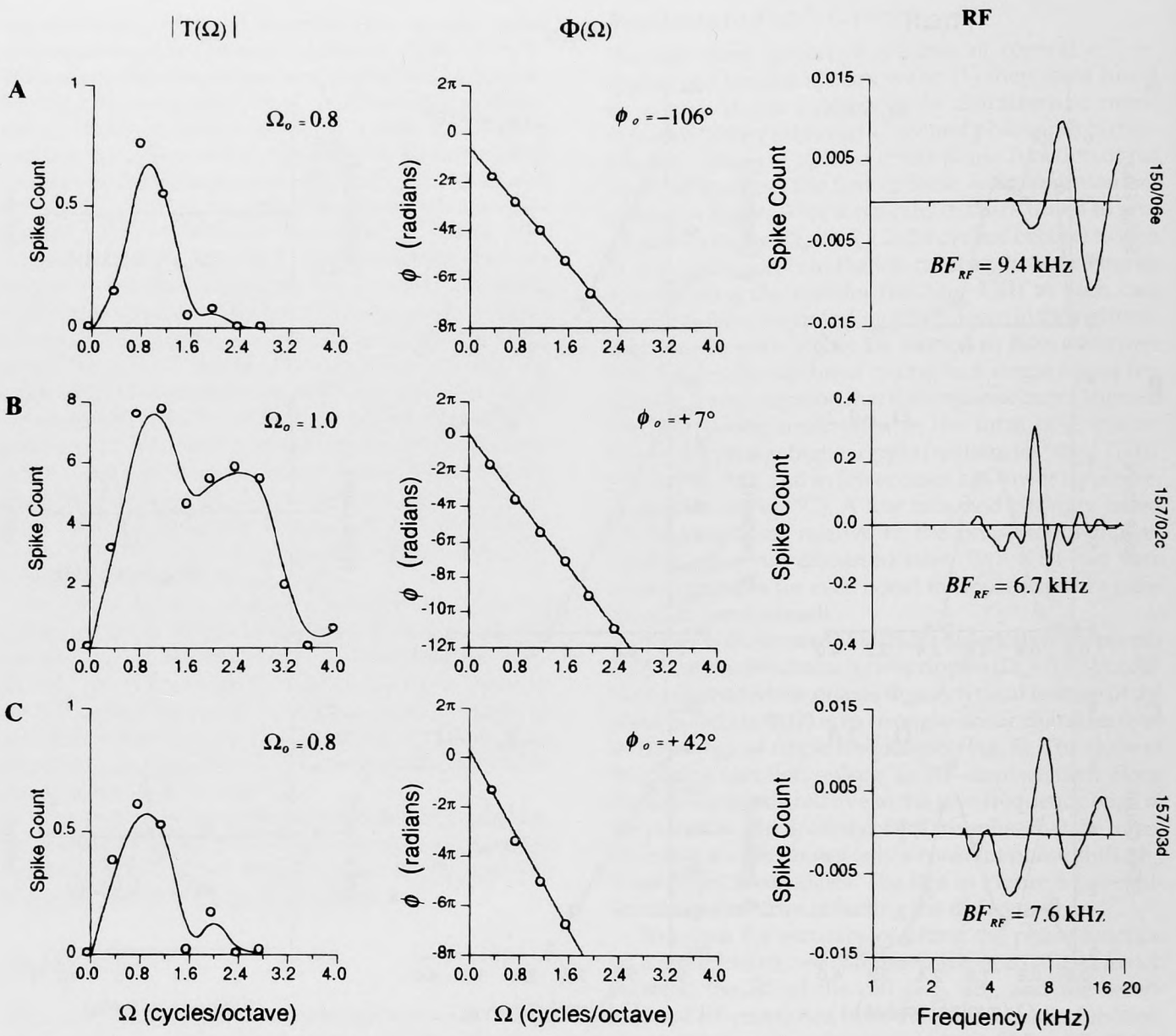


FIGURE 6 **Examples of ripple responses from three cells (*A, B, C*) with different characteristic phases ($\Phi_0 = -106°, 7°, 42°$, respectively), but similar characteristic ripples (Ω_0 = 0.8 – 1.0 cycles/octave).** Other details of the plots are as in Figure 5. The phase-fit indicators for the three cells are 0.98, 0.98, 0.93.

Dependence of the Responses on Ripple Amplitudes

Ripple spectra were presented at various ripple amplitudes to assess the stability of the responses. Figure 9 illustrates typical responses measured in 2 of 10 cells so tested. In both units the transfer function was measured at three ripple amplitudes. Over an intermediate range of amplitudes, the overall scale of the magnitude transfer function increases with ripple amplitude. At large amplitudes (such as at 50% in Fig. 9A), the growth of $|T(\Omega)|$ saturates but the overall shape of the transfer function is not much affected. With decreasing amplitude, the responses gradually weaken, and the $|T(\Omega)|$ becomes more noisy before disappearing (threshold at about 1 dB, or 25%). For cells with a high Ω_0 ($\geq$ 1 cycles/octave) Ω_0 slightly decreases with increasing ripple amplitude (t = 8; n = 5; P <0.01). Figure 9 demonstrates that the phase function $\Phi(\Omega)$ and the RF remain relatively stable with ripple amplitude. Tested over all 10 cells we found that the parameters Φ_0, BF_{RF}, and the $|T(\Omega)|$ width did not vary significantly (paired t test comparing responses to small and large amplitude: $t < 1.0$; n = 10; p >0.1).

Dependence of the Responses on Stimulus Level

Response properties were examined with respect to a change in overall stimulus level in 35 cells. Typical re-

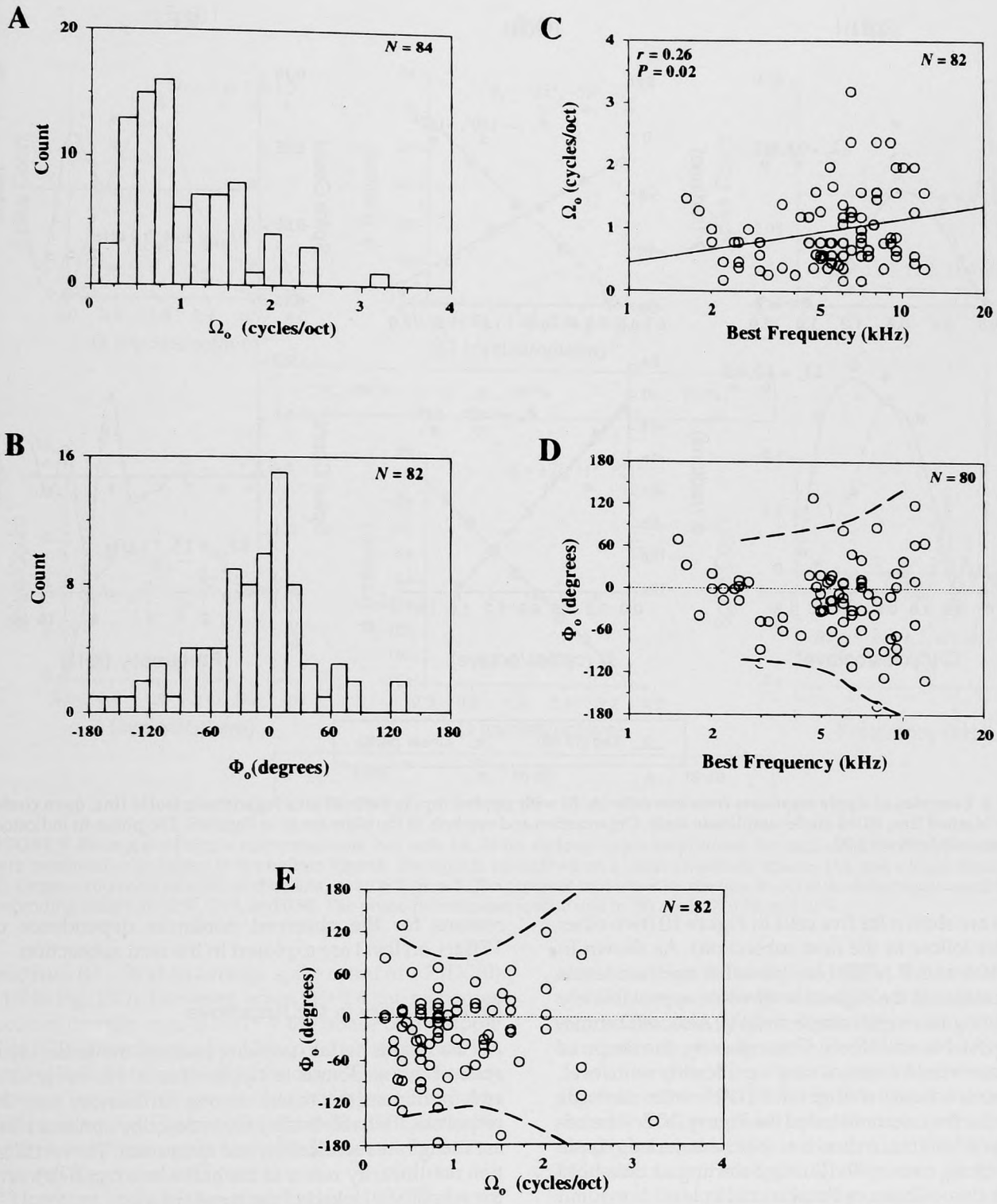


FIGURE 7 **Distributions of ripple response parameters in single-unit recordings in AI.** (A) Distribution of characteristic ripple Ω_0 (B) Distribution of characteristic phase Φ_0. (C) Distribution of Ω_0 as a function of BF. The solid line represents the linear regression, which indicates a weak but significant correlation ($r = 0.26$, $P < 0.05$). (D) Distribution of Φ_0 as a function of BF. (E) Combined distribution of Ω_0 and Φ_0. At four intervals of Ω_0 (0–0.6; 0.6–1.2; 1.2–1.8; > 1.8) the means and SD of Φ_0 were computed; the dashed lines represent a smoothed connection of the 2 SD edges.

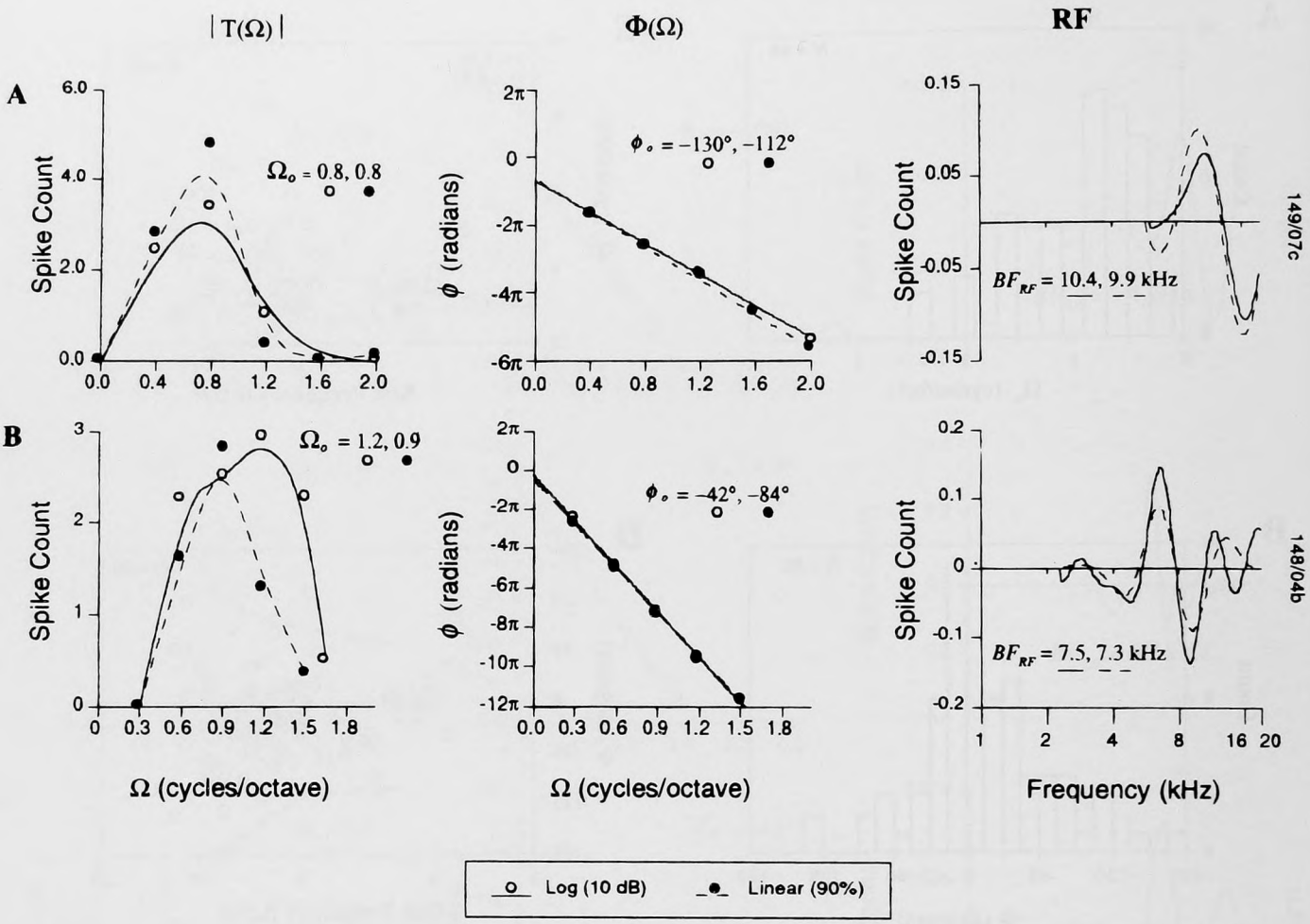


FIGURE 8 **Examples of ripple responses from two cells (A, B) with rippled inputs defined on a logarithmic (solid line, open circle) or linear (dashed line, filled circle) amplitude scale.** Organization and symbols of the plots are as in Figure 5. The phase-fit indicators for the two cells both are 1.00.

sponses are shown for five cells in Figure 10 (two other examples follow in the next subsection). As shown in Figure 10A and B $|T(\Omega)|$ is optimal at medium levels and decreases at the highest level while approximately maintaining its overall shape and Ω_0. Also, $\Phi(\Omega)$ functions are stable with level. Consequently, the shape of the response field does not vary significantly with level. The nonmonotonic change of $|T(\Omega)|$ with stimulus level is further demonstrated in Figure 10C. The responses of 3 units are shown at the characteristic ripple frequency Ω_0 over a 30 dB range starting at threshold level. In two cells, an optimal stimulus level is evident within the 30 dB range. In the other, the responses grew slower with level and whether $|T(\Omega)|$ decreased at higher stimulus levels (more than 65 dB) could not be confirmed. In 12 of 18 cells in which we measured curves as in Figure 10C, nonmonoticity was demonstrated. The nonlinear dependence of $T(\Omega)$ on level occurred both in units that had nonmonotonic (Fig. 10A) or monotonic rate-level functions for BF tones (Fig. 10B). The possible reasons for the observed nonlinear dependence of $|T(\Omega)|$ on level are explored in the next subsection.

Nonlinearities in the Responses

All the results so far have emphasized the linear character of the responses to ripple stimuli. However, several nonlinearities exert strong influences on the responses. Two relatively easy to describe nonlinearities are firing rate rectification and saturation. The rectification nonlinearity refers to the half-wave rectification of the responses locked to the ripple phase.

Figures 11 and 12 show the effects of nonlinearities on the ripple responses for cells with nonmonotonic and monotonic rate-level functions at BF tones (plots *A*), respectively. The effects of the rectification nonlinearity are most evident at the lower stimulus levels, that is, for the 50 dB responses in both examples shown. The effects are similar for nonmonotonic and monotonic cells. For instance, the nonmonotonic cell (Fig. 11) responds to a 50 dB flat

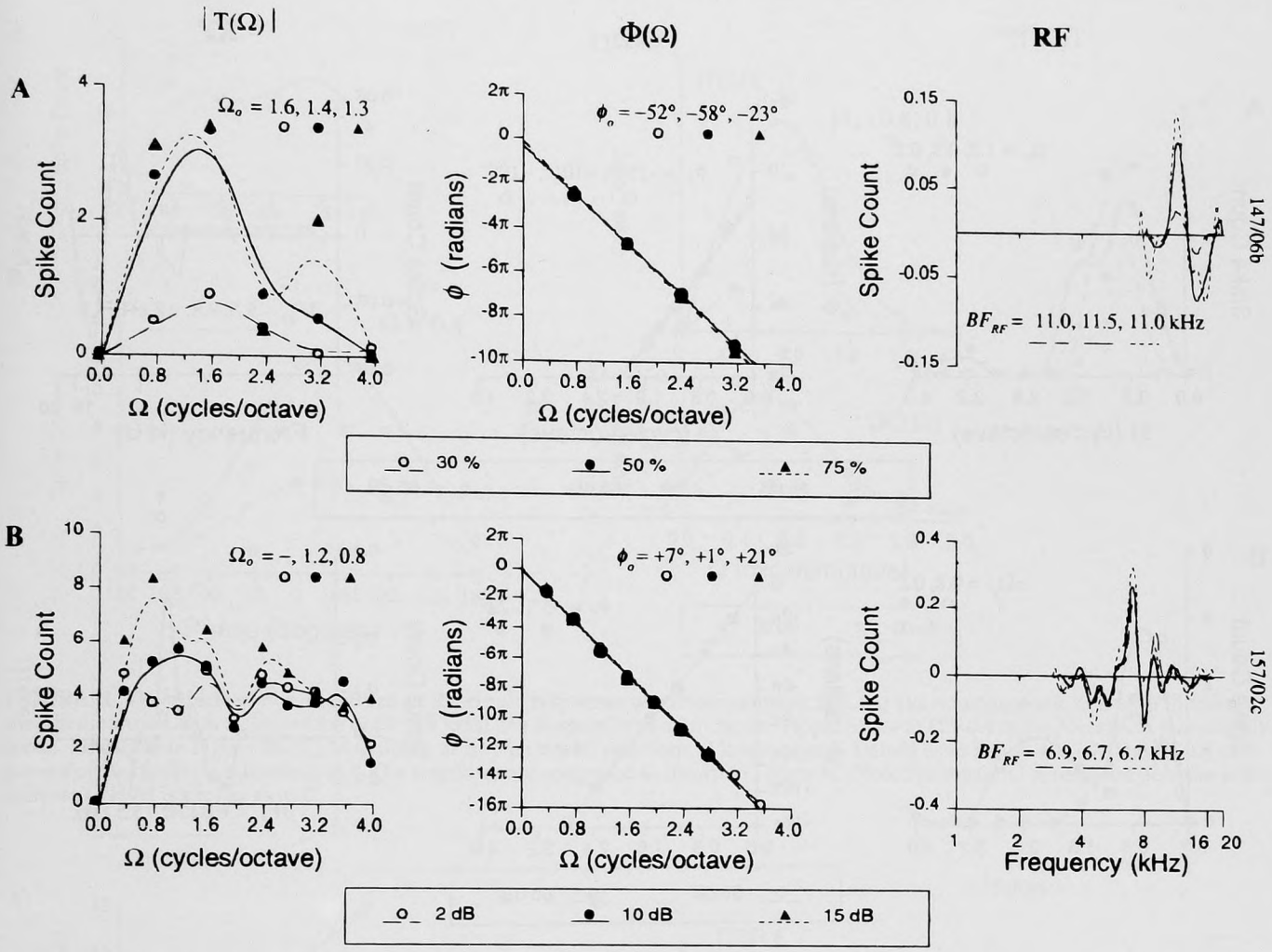


***FIGURE 9* Examples of ripple responses from two cells (A, B) for various ripple amplitudes.** For both cases, three ripple amplitudes were presented as indicated in the bottom legend. The ripples are defined on a linear amplitude scale in (A), and a logarithmic scale in (B). Organization and symbols of the plots are as in Figure 5. The phase-fit indicators for the unit in (A) at the three ripple amplitudes (in descending order) are 0.99, 0.99, and 0.98. The phase-fit indicators for the unit in (B) are 0.87, 0.89, and 0.74.

spectrum ($\Omega = 0$) at an average spike count of 1.7 ($DC(0) = 1.7$ in Fig. 11D). However, when $\Omega = 0.4$, the responses become strongly modulated by the phase of the ripple (Fig. 11B), increasing up to 12 in one-half cycle, and becoming half-wave rectified in the other. This rectification creates a large $DC(\Omega)$ component that is proportional to $|T(\Omega)|$, as seen in Figure 11D (plot at 50 dB). Exactly the same observations apply to the 50 dB data of the monotonic cell illustrated in Figure 12.

However, at the higher stimulus level (60 dB), the responses of the two types of cells diverge significantly. For the nonmonotonic cell, all phase-locked responses become suppressed (Fig. 11B), and hence both $|T(\Omega)|$ and $DC(\Omega)$ decrease together in amplitude (Figs. 11C and D at 60 dB). For the monotonic cell in Fig. 12, increasing the stimulus level to 60 dB drives the responses harder into saturation. Consequently, the $DC(\Omega)$ response increases with level (Fig. 12D). In contrast, the phase-modulated responses (Figs. 12B and C) decrease, and hence $|T(\Omega)|$ and $DC(\Omega)$ are inversely related. Note also that $|T(\Omega)|$ is most distorted (saturated) near Ω_o, causing the entire function to appear slightly shifted downward (Fig. 12C).

In a few cells, the response patterns exhibited complex features that could not be fit within the data analysis framework presented so far. For instance, 4 units had a strongly tuned $DC(\Omega)$, and only weak $T(\Omega)$ responses. In 3 others, the tuned $DC(\Omega)$ responses could not be simply related to $T(\Omega)$, as was the case earlier in Figures 11 and 12. Such a response is illustrated in Figure 13, where the $T(\Omega)$ is tuned to a rather low $\Omega_o = 0.4$, whereas $DC(\Omega)$ is strongly tuned to a much higher and apparently unrelated ripple frequency (5 cycles/octave). Since most cells in our sample were tested over a limited ripple frequency range (0 to 4 cycles/octave), it is unclear whether the $DC(\Omega)$ tuning at high ripple frequencies is common.

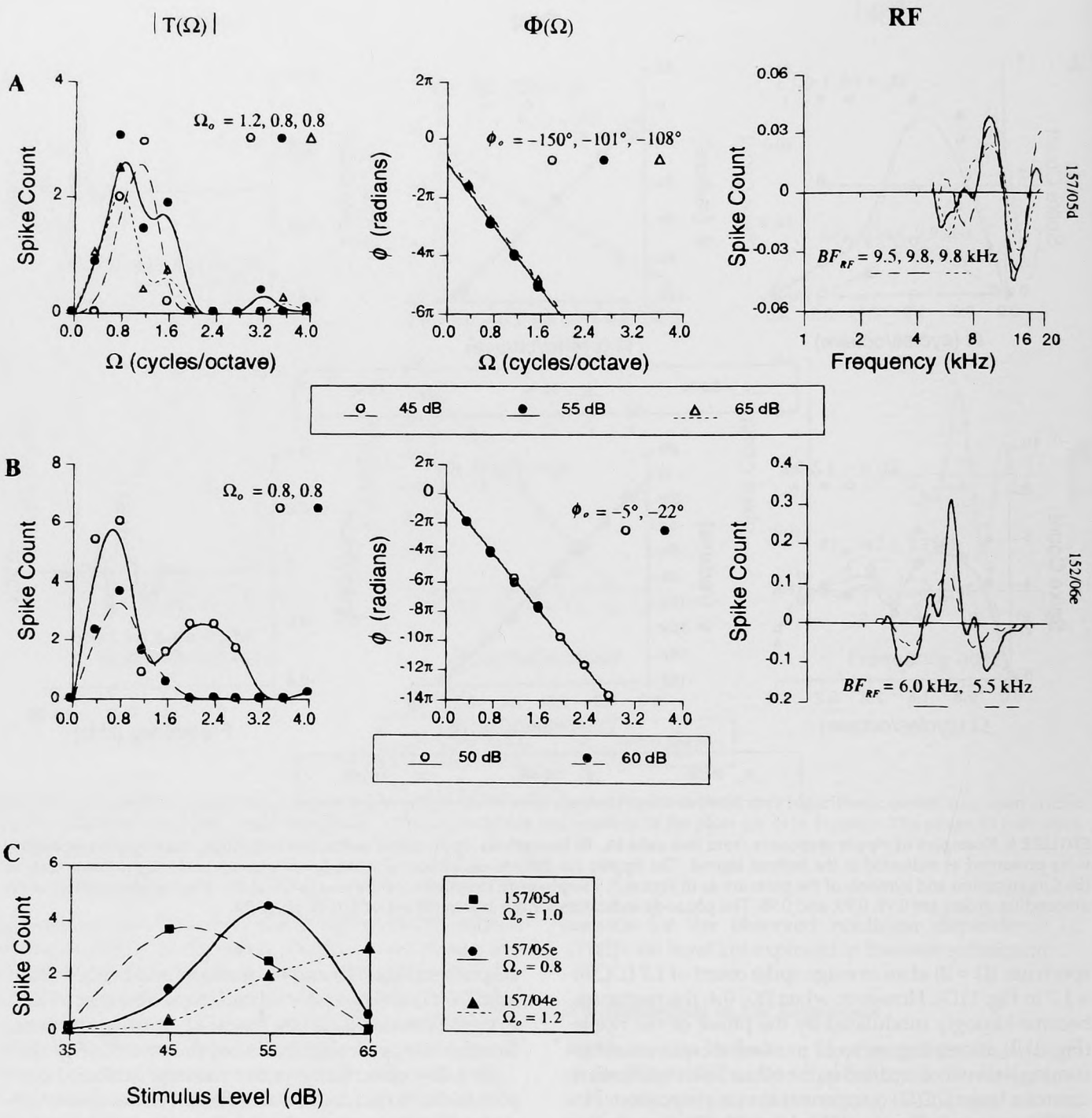


FIGURE 10 **Examples of ripple responses as a function of overall level of the stimulus.** Plots in *A* and *B* are organized as in Figure 5. (A) A nonmonotonic unit tested at three stimulus levels. (B) A monotonic unit tested at two stimulus levels. (C) Response magnitude at the characteristic ripple, $|T(\Omega = \Omega_0)|$ as a function of overall stimulus level for three cells.

Comparison Between the Response Field and Response Area

Figure 14A compares RFs (solid lines) to response areas as derived from two-tone responses (dashed lines) for three different cells. The RFs of different asymmetries and bandwidths match their response area counterparts (apart from the artificial elevation of responses toward the response area edges due to the BF tone (T2) of the two-tone stimulus). A qualitatively good correspondence is found in most cells, even in cells that have atypical RFs and response areas. An example of such responses is a double-peaked excitatory tuning, which is illustrated in Fig. 14B.

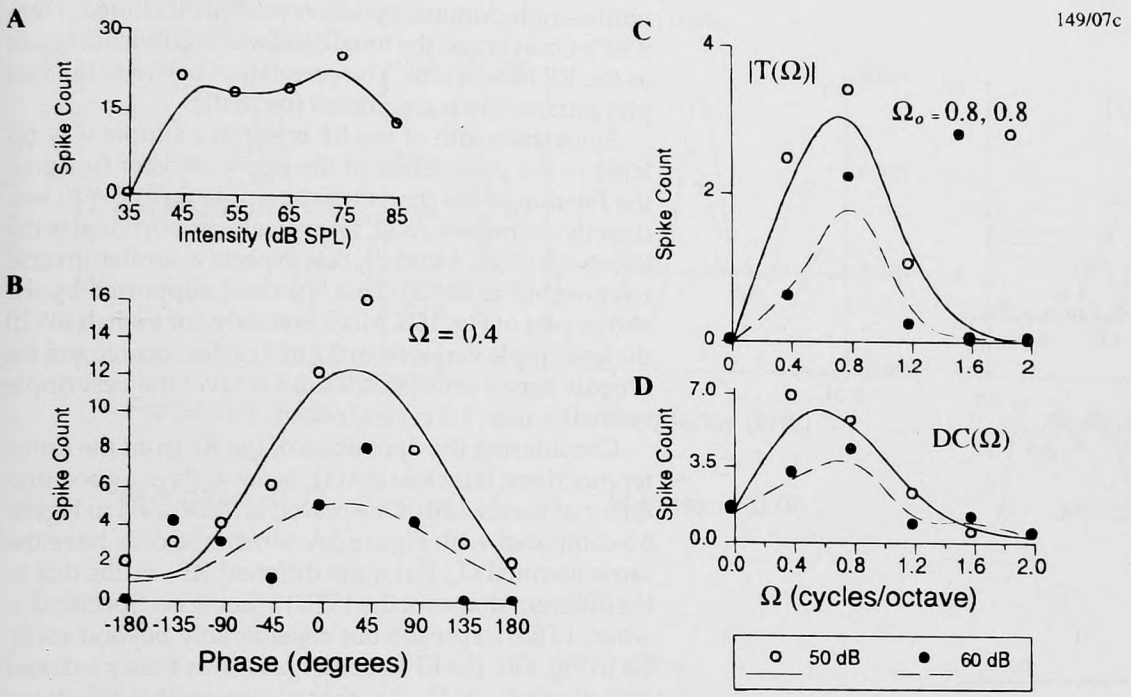


FIGURE 11 **The effects of nonlinearities on the ripple responses of a nonmonotonic cell.** (*A*) The nonmonotonic rate-level function of the cell measured with a single tone at BF (8.5 kHz). (*B*) Responses as a function of ripple phase at Ω = 0.4 cycles/octave, at two overall levels. Details as in Figure 3B. (*C*) Magnitude of ripple transfer functions at levels as in *B*. Details as in Figure 4A. (*D*) The *DC*(Ω) component of the response at levels as in *B*. The responses are computed as shown in Figure 3C. Note that the trend of response decrease with increase of level is similar as in *C*.

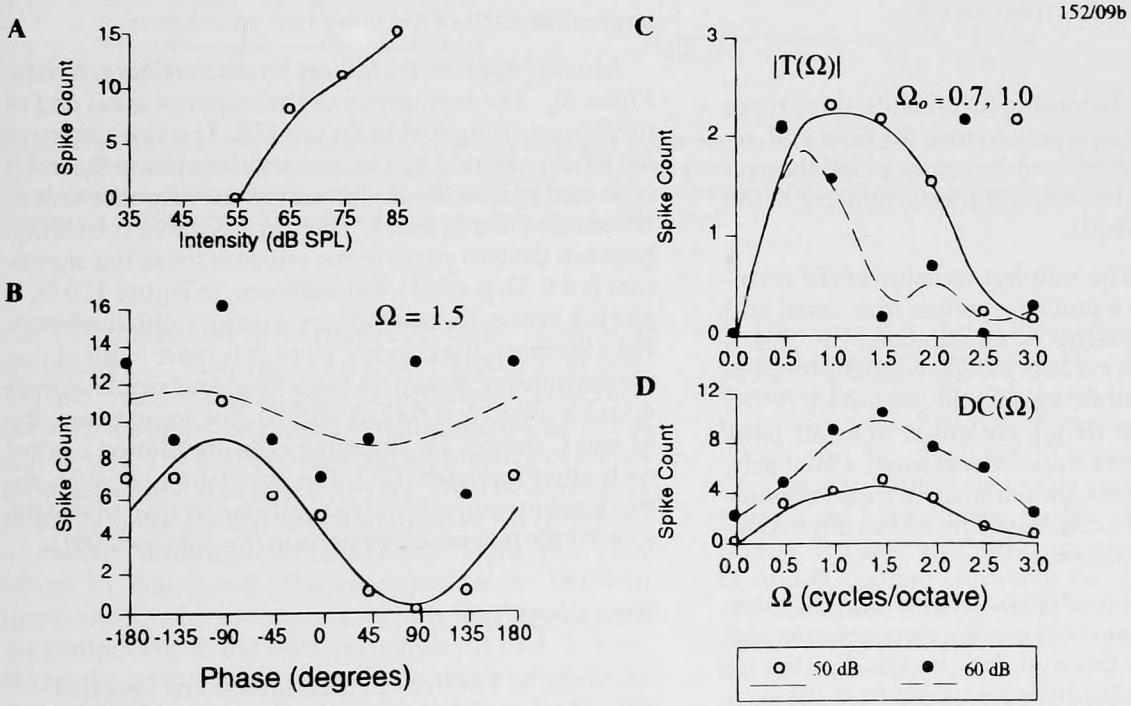


FIGURE 12 **The effects of the nonlinearities on the ripple responses of a monotonic cell.** All plots as in Figure 11. Plots *B*, *C*, and *D* for two different overall stimulus levels. (*A*) The rate-level function at a BF tone of 2.0 kHz. (*B*) The phase-locked responses at Ω = 1.5 cycles/octave. (*C*) Magnitude of ripple transfer functions. (*D*) The *DC*(Ω) response. Note that the trends of *AC*(Ω) and *DC*(Ω) with level are opposite.

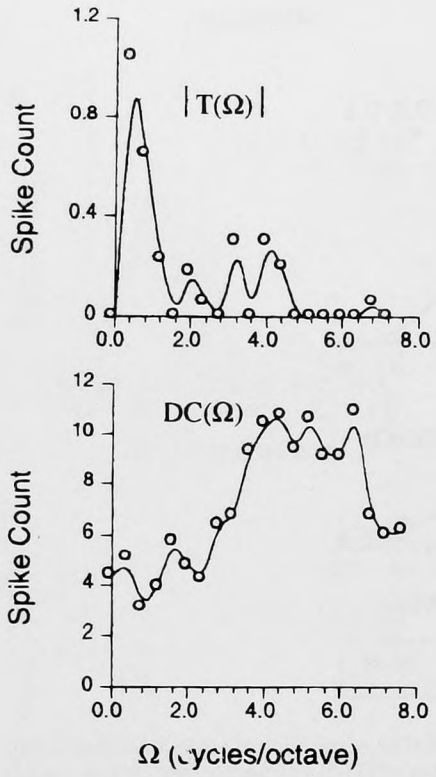


***FIGURE 13* Magnitude of ripple transfer function $|T(\Omega)|$ (*top*) and the $DC(\Omega)$ responses (*bottom*) of a more complex cell.** Plots as in Figure 11C and D, respectively. The two curves are uncorrelated to each other. The $DC(\Omega)$ is tuned at high ripple frequencies (4 to 7 cycles/octave), whereas the *AC* component is tuned to a low ripple frequency (0.4 cycles/octave).

To quantify this relationship for all units, three types of parameters were computed from the RFs and response areas, and juxtaposed in scatter plots; they reflect the BF, the bandwidth, and the asymmetry of the response areas and fields.

BF Versus BF_{RF}. The simplest measure of the correspondence between a unit's responses from tonal and rippled stimuli is in terms of its estimated BF. This is shown in Figure 15 as a scatter plot between the BF measured with a single tone against the frequency of the maximum of the RF (BF_{RF}), compiled from all units where both parameters could be measured. Obviously, for a majority of the cells, the parameters were very similar. Only for 9% of the cells the BF measures differed by more than half of an octave.

BW20 Versus Characteristic Ripple Ω_o. Another parametric comparison between the response area and the RF of a cell is in terms of their widths. For the response area, the bandwidth is measured from the excitatory tone responses 20 dB above threshold, BW20. For the RF, we measured the width of the positive peak between the zero crossings. Figure 16A shows the scatter plot of the RF width versus BW20. In this plot only RFs with a single dominant positive peak are included ($|\Phi_0| \leq 90°$). On average, the tonal bandwidth is twice as broad as the RF bandwidth. The correlation between the two plot parameters is significant ($p < 0.01$).

Since the width of the RF is not in a simple way related to the parameters of the ripple transfer function, the relation of the characteristic ripple Ω_0 to BW20 was directly examined. As Ω_0 is inversely proportional to the RF width (Figs. 4 and 5), one expects a similar inverse relationship to BW20. This is indeed supported by the scatter plot of Fig. 16B. More precisely, for a small BW20 the best ripple varies from 0.2 to 3 cycles/octave and for broadly tuned cells (BW20 > 1.5 octave) the best ripple is smaller than 1.2 cycles/octave.

Considering the derivation of the RF from the transfer functions, it is clear that Ω_0 is not always a good predictor of bandwidth. One reason is illustrated in Figure 6B compared with Figure 5A, where the cells have the same nominal Ω_0, but quite different RF widths due to the different shapes of the $|T(\Omega)|$ functions. Specifically, when $|T(\Omega)|$ spreads out considerably beyond its Ω_0 (as in Fig. 6B), the RF becomes narrower than predicted exclusively by its Ω_0. Another parameter that affects the bandwidth is Φ_0. This is demonstrated in Figure 4C where, for the same Ω_0, the RF width increases with about 50% as Φ_0 is changed from 0 to $\pi/2$. Multiple regression analysis confirmed that the RF width is inversely correlated to Ω_0 and the width of $|T(\Omega)|$ (measured at the 3 dB points), and positively correlated to $|\Phi_0|$. The regression weight of Ω_0 is about a factor 4 larger than each of the other two parameters.

M and C Asymmetry Indices Versus the Characteristic Phase Φ_0. The asymmetry of the response areas and of the RFs are compared in Figure 17A. The asymmetry of the RF is indicated by the characteristic phase Φ_0 and it is plotted against the asymmetry of the response area as reflected by the *M* index. There is a positive correlation between the two parameters, which is weak but significant ($r = 0.33$; $p < 0.01$). Furthermore, in Figure 17B Φ_0 is plotted versus the asymmetry measure obtained from FM responses, the *C* index. Here, it is more evident that the asymmetry measures from tonal and ripple stimuli do correspond ($r = 0.42$; $p < 0.001$). For comparison, the *M* and *C* indices for the same cells are plotted against each other in Figure 17C. The correlation between the two tonal measures is only slightly larger than that of the tone-ripple response comparison ($r = 0.46$; $p < 0.001$).

DISCUSSION

Linearity of Cortical Cell Responses to Spectral Ripples

Responses of single units in AI suggest that there is generally a good correspondence between the shape of a response area (measured with tonal stimuli) and that of

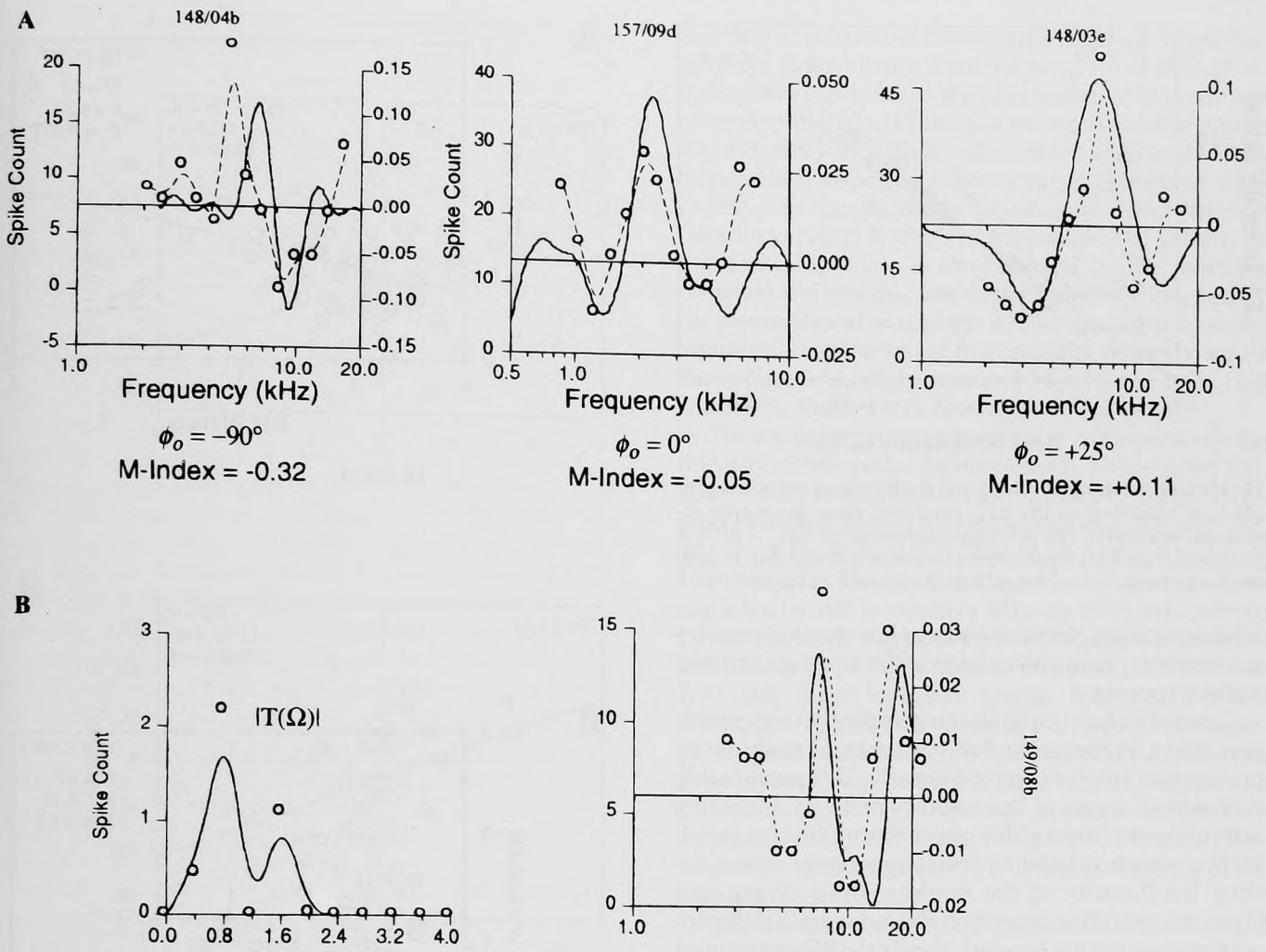


***FIGURE 14* Comparison between tonal response areas and ripple response fields (RFs).** (A) Examples of responses from three cells. The response areas (dashed lines) are measured using the two-tone paradigm, and hence the response counts are artificially elevated toward the edges (Shamma et al., 1993). The RFs and response areas exhibit similar bandwidths, asymmetries, and BFs. (B) An atypical example of a cell with both RF and excitatory response area double tuned.

the RF (measured with ripple transfer functions). Specifically, for most cells BF_{RF} is very similar to BF (Fig. 15), and there is a small but still significant correlation between the response area bandwidth and the RF width and Ω_0 (Fig. 16), and between the asymmetry of the response area and that of the RF (Fig. 17). These findings in single units are confirmed in another set of experiments with multiunit recordings (see Versnel *et al.* elsewhere in this issue). The correspondences between ripple and tonal responses must imply that there is a linear component in the cells' responses, or that, at least, the nonlinearities do not completely distort such linear response measures as the transfer function $T(\Omega)$. This conclusion is consistent with findings by Nelken *et al.* (1994) that responses to complex sounds (such as nine-tone complexes) are predictable on the basis of the response areas measured with two-tone stimuli.

However, it is also evident from the various scatter plots that, apart from the BF-BF_{RF} comparison, the correlations between the tonal and ripple response parameters are weak. Beside various response nonlinearities, which will be discussed in more detail later, other sources may contribute to this scatter. Primary among them is the approximate nature of the response parameters. For example, although parameters such as the Ω_0 and Φ_0 capture efficiently the shape of the RF, they are incomplete descriptors in that they can vary over a certain range without causing significant distortion of the RF (as in all cases of Figs. 8, 9, and 10). Also, measurements of these and tonal parameters as BW20 and *M* index were performed at a limited range of stimulus settings (such as resolution of Ω, Φ, frequencies). Nevertheless, given these and other possible sources of measurement errors, the persistence of a significant cor-

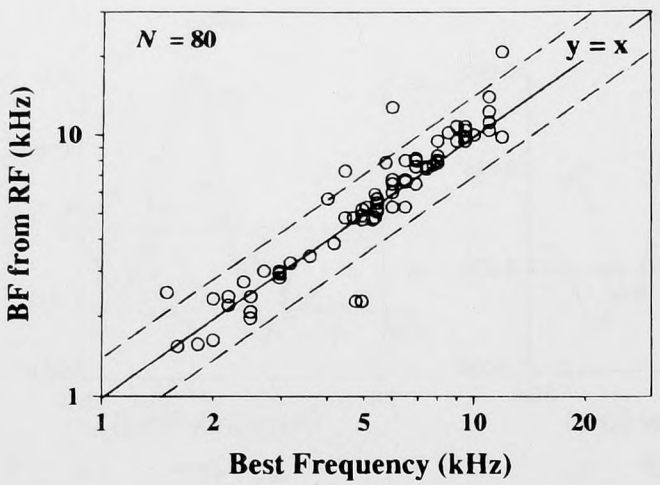


FIGURE 15 **Scatter plot comparing the BF measured with a single tone (abscissa) to the BF_{RF} predicted from the ripple responses (ordinate).** The solid line represents the BF_{RF} = BF line and the dashed lines represent the half-octave deviations. In only seven cases the deviations are half of an octave or larger.

relation between the tonal and ripple response measures is a testimony to the robustness of the linear component in the responses.

Linearity of cortical cells must imply that superposition holds, in the sense that responses to multiple simultaneous ripples combine linearly. This presumably is true at all stages of the auditory pathway, including at the output of the cochlea where the multitude of well-known peripheral nonlinearities apparently do not destroy the linearity of the ripple analysis (Wang and Shamma, 1994). Superposition also suggests that the ripple transfer function (or equivalently the RF) can be used to predict the responses of a cortical cell to any arbitrary spectral profile, since the profile can be thought of as a combination of multiple ripple components of various amplitudes and phases. This possibility is examined directly in Shamma and Versnel elsewhere in this issue.

Finally, it can be argued that ripple response measures (such as $T(\Omega)$) are potentially superior descriptors of cortical cell responses compared to those obtained with tones, although the two are theoretically equivalent. For instance, it is not possible to use the traditionally defined response area of a cell (as in Fig. 5 in Shamma *et al.*, 1993) to predict quantitatively its responses to an arbitrary spectrum. It is also experimentally difficult to measure accurately the RF for a cortical cell with a single tone because of the lack of vigorous spontaneous activity against which inhibition or subthreshold excitation can be quantified. Using two-tones to reveal these influences as two-tone inhibition (Suga 1977; Shamma *et al.*, 1993) or two-tone facilitation (Suga *et al.*, 1979; Nelken *et al.*, 1994) introduces other distortions in the RF, such as elevated background rates (Fig. 14) and response variability due to the relative tone intensities. In contrast, using ripples, the RF is a relatively stable and easily obtainable measure.

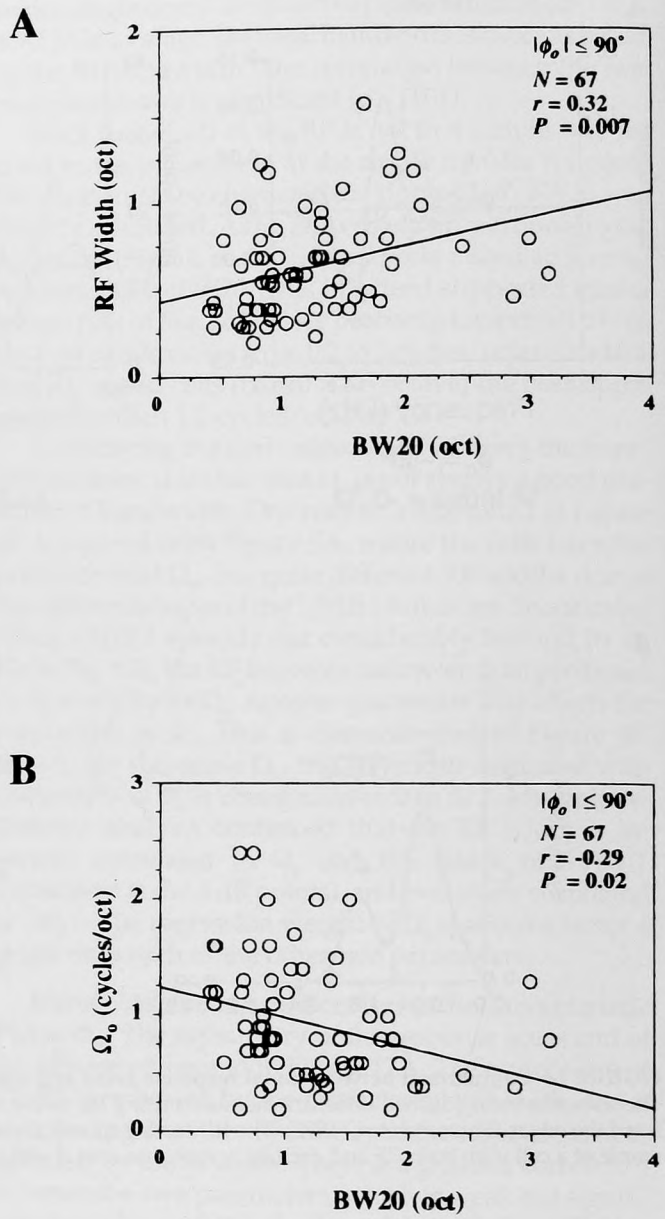


FIGURE 16 **Ripple response measures compared to the bandwidth of the response areas.** The solid lines represent linear regression lines. Correlation measures are indicated in the right-hand corner. In both plots only cells with $|\Phi_o| \leq 90°$ are used. (A) The width of the RF versus the bandwidth of the response area, BW20. (B) The characteristic ripple Ω_o versus BW20.

Functional Significance of the Response Characteristics

Aside from the linear component of their responses, the vast majority of units encountered in AI exhibited tuned responses as a function of ripple frequency. Furthermore, the characteristic ripples spanned a range of frequencies (0.2 to 3 cycles/octave), and were not simply clustered around one value (Fig. 7A). These findings suggest that AI cells can, in principle, function as ripple bandpass filters, analyzing an input spectral profile into separate

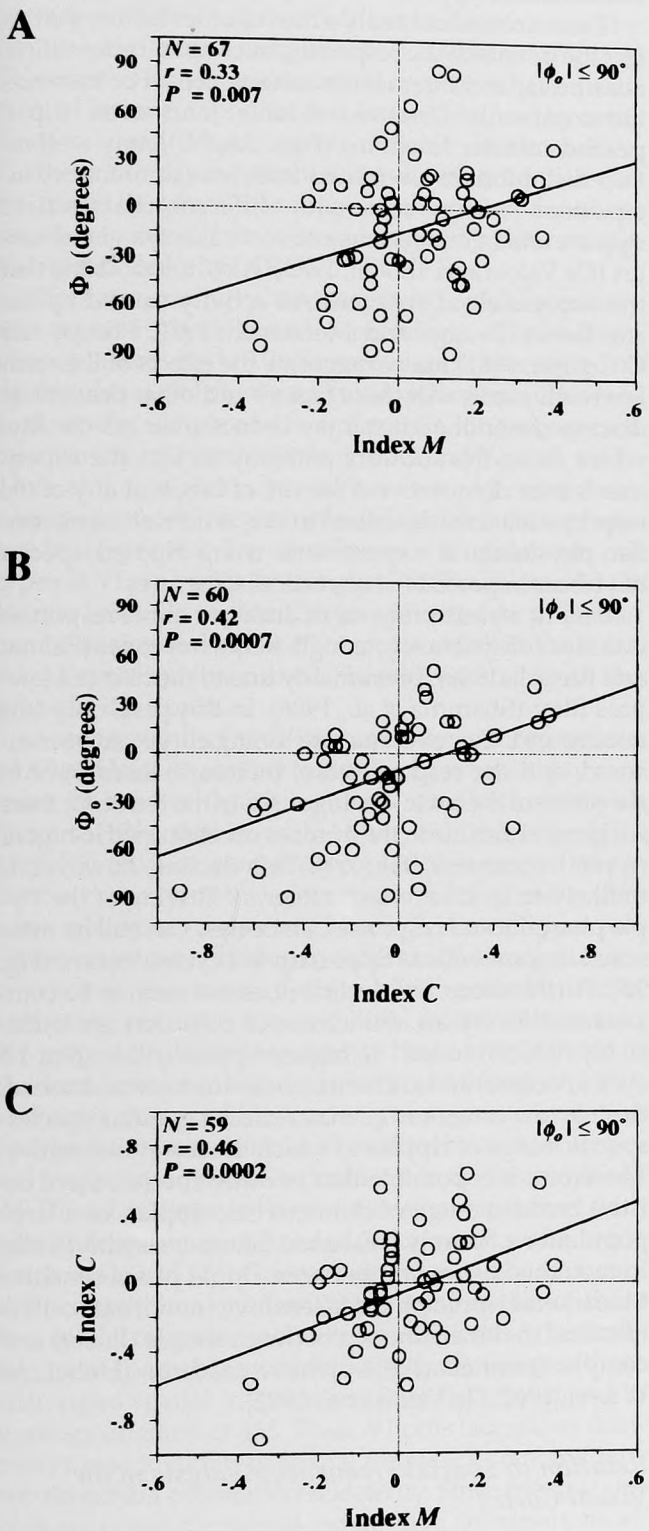


FIGURE 17 **The asymmetry of the RF reflected by characteristic ripple phase compared to tonal measures.** Only cells with $|\Phi_0| \leq 90°$ are used. (A) The characteristic phase Φ_0 versus the response area asymmetry, *M* index. (B) The characteristic phase Φ_0 versus FM direction sensitivity, *C* index. (C) Scatter plot between *C* index and *M* index.

channels tuned around different characteristic ripple frequencies. Equivalently, from the perspective of their response area bandwidths, they can be said to have a range of bandwidths so as to analyze the input spectral profile into different scales. Thus, cells with broader bandwidths respond best to the gross (slowly varying) features of the profile, whereas narrowly tuned units detect the fine (rapidly varying) features of the profile. The validity of this functional view is strengthened by the psychoacoustical findings discussed later in this section, and by the physiological mappings demonstrating a spatially organized distribution of this response property across the surface of AI (Schreiner and Mendelson, 1990; Heil *et al.*, 1992; Versnel *et al.* elsewhere in this issue).

The second important property of AI responses is the linearity of their phase functions $\Phi(\Omega)$. This implies that, apart from a linear phase shift due to the RF location relative to the left edge of the ripple stimulus (compare Fig. 4*B*), the transfer function $T(\Omega)$ of an AI cell has a constant phase Φ_0. The functional interpretation of this finding is that a unit is selective not only to a characteristic ripple frequency Ω_0, but also to a particular (characteristic) phase Φ_0 of that ripple. In this sense, AI cells analyze the input spectral profile into yet another dimension, namely, the phase of the ripples. This interpretation is consistent with findings that AI cells with asymmetrical response areas are selectively responsive to spectral profiles with the opposite asymmetry (Shamma *et al.*, 1993; Vranić *et al.*, 1993), since the characteristic phase is correlated with the asymmetry of the response area (Fig. 17).

There are, however, two important constraints on the ripple analysis model just presented. The first is that the distributions of the Ω_0 and Φ_0 are not uniform. Instead, Ω_0 in the ferret is largely limited below 2 cycles/octave, whereas Φ_0 is dominant around 0 (Fig. 7*B*). The second constraint is that each cell is only optimally responsive over a narrow range of stimulus levels (roughly, 20 dB) (Fig. 10*C*). This limitation, however, can easily be overcome if a population of cells are responsive to different "best" overall stimulus levels (as suggested by Fig. 10*C*).

In summary, AI response properties to rippled spectra suggest that it may function as a bank of ripple bandpass filters computing the local Fourier transform of the input spectral profile. A unit with a characteristic ripple frequency Ω_0 and phase Φ_0 would analyze the local region of the profile around its BF. The ripple-frequency axis can therefore be considered as a scale axis, with coarser views of the profile available at the low ripple filters, and finer details at the high ripple filters. Similarly, the ripple phase sensitivity axis can be seen as explicitly encoding the local asymmetry of the profile at each scale. The strength of its output reflects both the ripple frequency content near Ω_0, and the local asymmetry (relative to Φ_0) of that region of the profile. An array of such analyzers at different BFs, Ω_0s and Φ_0s would then perform the complete profile transformation.

Response Nonlinearities and Their Implications

The correspondence observed between the various response measures using tonal and rippled stimuli suggests that rectification and saturation nonlinearities do not completely disrupt the linear character of the responses. Linear response measures such as the ripple transfer function may indeed provide a meaningful characterization of a unit's response area. Perhaps the best way to explain this apparent paradox is by analogy to the effects of cochlear nonlinearities on the phase-locked responses of the auditory nerve.

The threshold nonlinearity of the cochlear hair cell rectifies the phase-locked responses of the auditory-nerve creating DC and higher harmonics that follow very similar trends as those described earlier in conjunction with Figures 11 and 12 (Dallos and Santos-Sacchi, 1983; Shamma *et al.*, 1986). Thus, it is possible to measure on an auditory nerve fiber an AC transfer function [analogous to $T(\Omega)$ in AI] based entirely on the phase-locked responses of the fiber using a swept single tone (Rose *et al.*, 1971) or reverse-correlation techniques with a noise stimulus (De Boer and De Jongh, 1978). Because of hair cell threshold, the synchronous responses are normally accompanied by a similarly tuned DC component (usually called the "average rate" response), much like the $DC(\Omega)$ component in AI responses (see Shamma *et al.*, 1986 for a detailed discussion of these different response measures).

The saturation nonlinearity limits the growth of the ripple phase-locked responses in AI cells (Fig. 10) in an analogous manner to that seen in the synchronous responses on the auditory nerve. Note, however, that the decrease of $|T(\Omega)|$ at the highest stimulus levels is not analogously seen in auditory nerve fibers (Rose et al., 1971) because they normally operate near the threshold knee of the nonlinear transfer characteristics of the hair cell and not near saturation.

Despite rectification and saturation nonlinearities, AC transfer functions on the auditory nerve are a valuable predictive measure of a fiber's response to broadband stimuli (Deng *et al.*, 1988). It is in this sense that one may conjecture that ripple transfer functions $T(\Omega)$ are also useful in describing AI responses to a spectral profile.

A different source of nonlinear interactions is the uncertain "internal" representation of a spectral ripple at the input of the central auditory system. Both linearly and logarithmically amplitude modulated rippled spectra were used as spectral profiles in our experiments. Clearly, one or both of these sinusoidal spectra may appear distorted to the central auditory system. This kind of "input nonlinearity" creates additional DC and higher harmonic components that are not explicitly accounted for in the stimulus. Nevertheless, the experimental results summarized in Figure 8 suggest that the two inputs produce similar responses, suggesting that the distortion harmonics are smaller than the primary ripple and, for each unit, lie sufficiently away from its characteristic ripple (that is, outside of the $|T(\Omega)|$).

There are undoubtedly a host of other factors that affect the details of the responses, such as cell adaptation, anesthesia, and interactions among cells. For instance, the commonly observed double (and even triple) peaked transfer functions (Figs. 5A, 6B), may well reflect disinhibition associated with lateral inhibitory interactions among cells with different characteristic ripples, analogous to those observed in the visual cortex (De Valois and Tootell, 1983). Also, it is possible that the depression of spontaneous activity caused by the anesthesia (Brugge and Merzenich, 1973; Pfingst and O'Connor, 1981) may exaggerate the effects of the rectifier nonlinearity. All these factors and other previously discussed nonlinearities may induce their effects anywhere along the auditory pathway. In fact, the experiments here do not reveal the site of origin of any of the response features described in this article, since no similar physiological experiments using rippled spectra have been reported in subcortical structures.

Finally, a basic property of auditory nerve responses is the loss of synchrony to high-frequency tones (Palmer and Russell, 1986), presumably due to the hair cell low-pass filter (Shamma *et al.*, 1986). In this case, only (the nonlinear) DC or average rate tuning curve can be measured, and the responses are, therefore, insensitive to the phase of the tone. Analogously, in the ferret AI, there is a general decline in the number of cells tuned to higher ripple frequencies (Fig. 7*A*). This decline, however, is unlikely to be due to an analogous filtering of the ripple phase-locked responses, since they can still be measured in some cells at ripples up to 4 cycles/octave (Fig. 9*B*). Furthermore, the decline does not seem to be compensated for by an abundance of cells that are exclusively "$DC(\Omega)$ tuned" to higher ripples (more than 1.5 cycles/octave) and are hence phase-insensitive. Instead, this narrow range of Ω_0s may reflect a genuine species-specific range of ripples to which the ferret is sensitive. Therefore, it is possible that in other species, there exists a broader range of characteristic ripples, or a large population of purely DC-tuned, phase-insensitive cells. In fact, this distinction between ripple phase-sensitive (linear) and ripple phase-insensitive (nonlinear) cells is identical to the distinction between simple (linear) and complex (nonlinear) cells of the visual cortex (Hubel and Wiesel, 1962; De Valois *et al.* 1982).

Relation to Spatial Frequency Analysis in the Visual Cortex

Physiological and psychophysical experiments with ripple-like stimuli (or gratings) in the visual system have been carried out for more than two decades (see De Valois and De Valois [1988] for a thorough review). In VI, cells display similar responses to those described here in AI. For instance, the transfer function of a VI cell

is tuned around a specific grating frequency (usually called "spatial frequency"), and its inverse transform predicts well the receptive field of the cell measured by impulse-like stimuli as light dots (De Valois *et al.* 1982). Thus, just as in AI, visual cortical responses have a substantial linear component that is not disrupted by rectification, saturation, and other nonlinearities. Furthermore, psychophysical data have accumulated in the visual literature that supports the notion that these tuned response features are perceptually relevant (De Valois and De Valois, 1988).

Ripple (or grating) phase has played only a secondary role in most physiological experiments of spatial frequency selectivity in VI. Nevertheless, it has served to distinguish cleanly between two fundamental classes of cortical cells, simple and complex cells, long recognized by a variety of other criteria (Hubel and Wiesel, 1962). Thus, although both cell types have tuned transfer function magnitudes, only simple cells exhibit clear phase-locked responses. The distributions of these two cell types in VI are basically comparable (De Valois and De Valois, 1988). By contrast, there are very few units that can be analogously called "complex cells" in the ferret AI, as most exhibit robust "linear" sensitivity to the phase of the ripple.

Another significant difference between the responses of AI and VI with respect to ripple phase sensitivity concerns the relative dearth in our AI sample of units with "reverse" RFs, that is, with $|\phi_0| > 150$ (as in Fig. 14B). Such RFs in VI (known as off-center-on-surround) are at least as common as their counterparts. It is possible they were missed in AI due to a bias in the sampling procedures, or are concentrated in subareas of AI as yet undiscovered in the ferret. Cells exhibiting doubly tuned excitatory response areas, which presumably would have $\Phi_0 \approx 180°$, have been reported to be concentrated in the dorsal area of AI in the cat (Sutter and Schreiner, 1991).

Finally, it is possible to give a simple interpretation of orientation selectivity in VI within the context of spatial frequency analysis, and relate it to ripple analysis in AI. Visual gratings are two-dimensional in nature with spatial frequencies defined along two axes of a scene (such as Ω_x and Ω_y). Oriented gratings can be uniquely defined (within a quadrant) by a combination of these two spatial frequencies. For instance, vertical (horizontal) gratings are those with $\Omega_y = 0$ ($\Omega_x = 0$), whereas those with equal spatial frequencies ($\Omega_x = \Omega_y$) correspond to gratings oriented at 45°. Thus, VI cells (simple or complex) tuned to different spatial frequency combinations would exhibit orientation selectivity. Since spectral ripples are one-dimensional, orientation selectivity as already defined simply reduces to tuning along one dimension, that is, the usual characteristic ripple (Ω_0). However, it is crucial to recognize that apart from the dimensionality of the input signal, the mechanisms giving rise to orientation selectivity in VI are identical to those seen in AI.

Relationship to Psychoacoustics

We suggested so far that a spectral profile of an acoustic stimulus is analyzed by AI filters, each operating locally on the tonotopic axis and each filtering both ripple frequencies to encode the profile scale and ripple phases to encode the profile asymmetry. There are many psychoacoustical implications of this hypothesis, but very few experiments have been carried out so far to test them. One of the earliest studies by Green (1986), and a recent more detailed elaboration on it by Hillier (1991), tested the sensitivity of human subjects to rippled spectra of different frequencies. They both found increased sensitivity and uniformity of thresholds around approximately 1 to 3 cycles/octave; nevertheless, subjects responded over a broad range of ripples well exceeding 10 cycles/octave. Similarly, recent measurements (Vranić-Sowers and Shamma, 1995b) revealed that ripple phase sensitivity up to approximately 1 cycle/octave is constant at around 6° regardless of ripple frequency. Above 1 cycle/octave, threshold increased gradually. These threshold values and trends are almost identical to those found in corresponding visual experiments (De Valois and De Valois, 1988).

In summary, it is hypothesized that the auditory system analyzes locally the spectral profile along a ripple frequency (scale) and a phase (asymmetry) dimension (Vranić-Sowers and Shamma, 1995a). An implication of this hypothesis is that the perception of timbre, as far as determined by the shape of the spectral profile, may be more accurately described using the ripple analyzed profile, rather than the profile itself. Similarly, and more generally, descriptions of higher level auditory perceptual tasks involving complex sounds, such as spatial localization, pitch perception, speech recognition, and detection of spatial and spectral motion, should be based on an explicit representation of the ripple transformation of the spectral profile.

ACKNOWLEDGEMENT

This work is supported by grants from the Air Force Office of Scientific Research, and from the Office of Naval Research. We would like to 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 for his assistance in surgery and data recordings. The authors are members of the Institute for Systems Research, which is partially funded by National Science Foundation grant (NSFD CD 8803012).

REFERENCES

Bilsen, F. A., Ten Kate, J. H., Buunen, F., and Raatgever, J. (1975). Response of single units in the cochlear nucleus of the cat to cosine noise. *J. Acoust. Soc. Am. 58*, 858–866.

Brugge, J. F., and Merzenich, M. M. (1973). Responses of neurons in auditory cortex of the Macaque monkey to monaural and binaural stimulation. *J. Neurophysiol. 36*, 1138–1158.

Dallos, P., and Santos-Sacchi, J. (1983). AC receptor potentials from hair cells in the low-frequency region of the guinea pig cochlea. In: *Mechanisms of Hearing*, Webster, W. R., and Aitkin, L. M., eds., Monash University Press, Australia, pp. 11–16.

De Boer, E., and De Jongh, H. R. (1978). On cochlear encoding: Potentialities and limitations of the reverse-correlation technique. *J. Acoust. Soc. Am. 63*, 115–135.

Deng, L., Geisler, C. D., and Greenberg, S. (1988). A composite model of the auditory periphery for the processing of speech. *J. Phonetics 16*, 93.

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

De Valois, K. K., and Tootell, R. B. H. (1983). Spatial-frequency-specific inhibition in cat striate cortex cells. *J. Physiol. (Lond.) 336*, 359–376.

De Valois, R. L., Albrecht, D. G., and Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Res. 22*, 545–559.

Evans, E. F. (1979) Single-unit studies of mammalian cochlear nerve. In: *Auditory investigation: The technological and scientific basis*, H. A. Beagley, eds., Oxford: Clarendon Press, pp. 324–367.

Green, D. M. (1986). 'Frequency' and the detection of spectral shape change. In: *Auditory Frequency Selectivity*, Moore, B. C. J., and Patterson, R. D., eds., Plenum Press, Cambridge, pp. 351–359.

Greenwood, D. D. (1990). A cochlear frequency-position function for several species—29 years later. *J. Acoust. Soc. Am. 87*, 2592–2605.

Heil, P., Rajan, R., and Irvine, D. (1992). Sensitivity of neurons in cat primary auditory cortex to tones and frequency-modulated stimuli. II. Organization of response properties along the iso-frequency dimension. *Hear. Res. 63*, 135–156.

Hillier, D. (1991). Auditory processing of sinusoidal spectral envelopes. PhD thesis, Washington University, St. Louis.

Houtgast, T. (1977) Auditory-filter characteristics derived from direct-masking data and pulsation-threshold data with a rippled-noise masker. *J. Acoust. Soc. Am. 62*, 409–415.

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

Kelly, J. B., Judge, P. W., and Phillips, D. P. (1986). Representation of the cochlea in primary auditory cortex of the ferret *Mustela putorius*. *Hear. Res. 9*, 35–41.

Kiang, N. Y. S., Watanabe, T., Thomas, E. C., and Clark, L. F. (1965). Discharge patterns of single fibers in the cat's auditory nerve. M.I.T. Research Monograph No. 35. M.I.T. Press, Cambridge.

Mendelson, J. R., and Cynader, M. S. (1985). Sensitivity of cat primary auditory cortex (AI) neurons to the direction and rate of frequency modulation. *Brain Res. 327*, 331–335.

Nelken, I., Prut, Y., Vaadia, E., and Abeles, M. (1994). Population responses to multifrequency sounds in the cat auditory cortex: One- and two-parameter families of sounds. *Hear. Res. 72*, 206–222.

Oppenheim, A. & Schafer, R. (1990). Digital Signal Processing. Prentice-Hall, New Jersey.

Palmer, A. R., and Russell, I. J. (1986). Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells. *Hear. Res. 24*, 1–15.

Pfeiffer, R. R., and Kim, D. O. (1972). Response patterns of cochlear nerve fibers to click stimuli: Descriptions for cat. *J. Acoust. Soc. Am. 52*, 1669–1677.

Pfingst, B. E., and O'Connor, T. A. (1981). Characteristics of neurons in auditory cortex of monkeys performing a simple auditory task. *J. Neurophysiol. 45*, 16–34.

Pickles, J. O. (1986) The neurophysiological basis of frequency selectivity. In: *Frequency Selectivity in Hearing*, B. C. J. Moore, eds., Acadamic Press, London, pp. 51–122.

Rose, J. E., Hind, J. E., Anderson, D. J., and Brugge, J. F. (1971). Some effects of stimulus intensity on response of auditory nerve fibers of the squirrel monkey. *J. Neurophysiol. 34*, 685–699.

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

Schreiner, C. E., and Mendelson, J. R. (1990). Functional topography of cat primary auditory cortex: Distribution of integrated excitation. *J. Neurophysiol. 64*, 1442–1459.

Shamma, S. A., Chadwick, R. S., Wilbur, W. J., Morrish, K. A., and Rinzel, J. (1986). A biophysical model of cochlear processing: Intensity dependence of pure tone responses. *J. Acoust. Soc. Am. 80*, 133–145.

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.

Suga, N. (1977). Amplitude spectrum representation in the Doppler-shifted-CF processing area of the auditory cortex of the mustache bat. *Science 196*, 64–67.

Suga, N., O'Neil, W. E., and Manabe, T. (1979). Harmonic-sensitive neurons in the auditory cortex of the mustache bat. *Science 203*, 270–274.

Summers, V., and Leek, M. R. (1994). The internal representation of spectral contrast in hearing-impaired listeners. *J. Acoust. Soc. Am. 95*, 3518–3528.

Sutter, M. L., and Schreiner, C. E. (1991). Physiology and topography of neurons with multipeaked tuning curves in cat primary auditory cortex. *J. Neurophysiol. 65*, 1207–1226.

Vranić, S., Versnel, H., and Shamma, S. A. (1993). Single and double spectral peaks: Psychoacoustical and physiological results. Association of Research Otolaryngology 16, St Petersburg Beach, Florida, Abstract 188.

Vranić Sowers, S., and Shamma, S. A. (1995a). Represen-tation of spectral profiles in the auditory system. I. A ripple analysis model. *J. Acoust. Soc. Am.* (in press).

Vranić Sowers, S., and Shamma, S. A. (1995b). Representation of spectral profiles in the auditory system. II. Detection of spectral peak shape changes. *J. Acoust. Soc. Am.* (in press).

Wang, K., and Shamma, S. A. (1994). Self-normalization and noise robustness in auditory representations. *IEEE Trans. Audio. Speech.1 Proc. 2*, 421–435.

Yost, W. A., Hill, R., and Perez-Falcon, T. (1978). Pitch and pitch discrimination in broadband signals with rippled power spectra. *J. Acoust. Soc. Am. 63*, 1166–1173.