JOURNAL OF NEUROPHYSIOLOGY
Vol. 76, No. 5, November 1996. *Printed in U.S.A.*

Analysis of Dynamic Spectra in Ferret Primary Auditory Cortex. II. Prediction of Unit Responses to Arbitrary Dynamic Spectra

NINA KOWALSKI, DIDIER A. DEPIREUX, AND SHIHAB A. SHAMMA
Institute for Systems Research and Electrical Engineering Department, University of Maryland, College Park, Maryland 20742-3311

SUMMARY AND CONCLUSIONS

1. Responses of single units and multiunit clusters were recorded in the ferret primary auditory cortex (AI) with the use of broadband complex dynamic spectra. Previous work has demonstrated that simpler spectra consisting of single moving ripples (i.e., sinusoidally modulated spectral profiles that travel at a constant velocity along the logarithmic frequency axis) could be used effectively to characterize the response fields and transfer functions of AI cells.

2. A complex dynamic spectral profile can be thought of as being the sum of moving ripple spectra. Such a decomposition can be computed from a two-dimensional spectrotemporal Fourier transform of the dynamic spectral profile with moving ripples as the basis function.

3. Therefore, if AI units were essentially linear, satisfying the superposition principle, then their responses to arbitrary dynamic spectra could be predicted from the responses to single moving ripples, i.e., from the units' response fields and transfer functions (spectral and temporal impulse response functions, respectively).

4. This conjecture was tested and confirmed with data from 293 combinations of moving ripples, involving complex spectra composed of up to 15 moving ripples of different ripple frequencies and velocities. For each case, response predictions based on the unit transfer functions were compared with measured responses. The correlation between predicted and measured responses was found to be consistently high (84% with $\rho > 0.6$).

5. The distribution of response parameters suggests that AI cells may encode the profile of a dynamic spectrum by performing a multiscale spectrotemporal decomposition of the dynamic spectral profile in a largely linear manner.

INTRODUCTION

Acoustic stimuli with broadband dynamic spectra evoke strong and relatively sustained responses in neurons of the primary auditory cortex (AI) (Eggermont 1994; Kowalski et al. 1996; de Ribaupierre et al. 1972). The response patterns reflect details of both the spectral shape and its changes in time. To characterize these neurons or units, elementary broadband spectra with envelopes sinusoidally modulated on a logarithmic axis (ripples) were presented over a wide range of parameters such as ripple frequencies, phases, and velocities (Kowalski et al. 1996). A typical and most prominent feature of the responses is the synchronized component that tracks the periodicity of the stimulus envelope. The amplitude and phase of this component could be measured from period histograms of the neural responses and plotted against different stimulus parameters, thus obtaining a variety of transfer functions. For example, the response component as a function of ripple frequency is the ripple transfer function, whose inverse Fourier transform is the response field of the unit ($\mathcal{RF}$). Similarly, the response as a function of ripple velocity gives the temporal transfer function and its inverse transform, the temporal impulse response ($\mathcal{IR}$).

Implicit in the use of $\mathcal{RF}$ and $\mathcal{IR}$ to describe AI responses is an assumption of linearity. That is, these "linear systems response measures" meaningfully characterize the response of a unit to the spectral shape on the one hand and its dynamics on the other. This is in fact the case for the $\mathcal{RF}$ in that AI responses to input spectra such as stationary ripples obey the superposition principle, and thus can be summed with appropriate weights to predict the responses to arbitrary combinations of stationary ripples (Shamma and Versnel 1995). This result also applies to moving ripples, because the $\mathcal{RF}$ shape (to within a scaling factor) is independent of the ripple velocities (Kowalski et al. 1996).

However, the relevance of the $\mathcal{IR}$ in characterizing AI unit responses is uncertain because the linearity of the temporal responses to dynamic spectra has not been directly demonstrated. The purpose of the experiments described here is to test directly whether the superposition principle holds for the responses elicited by moving ripples with different velocities. Specifically, our goal is first to record the responses of a unit to single ripples moving at a range of velocities. Then, we form the superposition of a specific combination of these responses (e.g., to a pair or triplet of moving ripples), and compare it with the actual responses obtained from a stimulus spectrum composed of the same ripple combination. If AI responses are linear, the two response patterns must be similar, and the $\mathcal{RF}$ and $\mathcal{IR}$ can be used to predict the responses of the unit to any arbitrary broadband dynamic spectrum.

METHODS

Surgery and animal preparation

Ferrets (*Mustela putorius*) were anesthetized with pentobarbital sodium (40 mg/kg) and maintained in an areflexic state by continuous intravenous infusion of pentobarbital and lactated Ringer solution mixed with dextrose. The ectosylvian gyrus, which includes AI, was exposed by craniotomy and the dura was reflected. The contralateral ear canal 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. For details on the surgery see Shamma et al. (1993).

Acoustic stimuli

Various auditory stimuli were used: pure tones (single tone bursts, 200-ms duration, 8-ms triangular rise and fall times), broad-

band complex sounds (single ripples), and linear combinations of these complex sounds (multiple ripples). These are briefly reviewed below; a more extensive description can be found in Kowalski et al. (1996). All complex stimuli were computer synthesized, gated, and then fed through an equalizer into the earphone. Calibration of the sound delivery system (up to 20 kHz) was performed in situ with the use of a 1/8-in. Bruël & Kjaer probe microphone (type 4170). The microphone was inserted into the meatus 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).

Ripples were made up of 101 tones equally spaced along a logarithmic frequency axis and spanning 4.32 or 5 octaves (e.g., 1–20 kHz, 0.5–16 kHz, or 0.25–8 kHz), such that the response area of the cell being tested lay within the stimulus' spectrum. The amplitude of each of the tones was chosen so that the spectral envelope of the resulting broadband stimulus forms a sinusoid (a ripple) on a linear amplitude scale, with the amplitude usually set at 90–100% modulation. Schematically, then, the envelope profile is given by

$$S(x) = 1 + \Delta A \cdot \sin(2\pi \cdot \Omega \cdot x + \Phi) \quad (1)$$

where ΔA is 0.9 or 1; x is the logarithmic frequency axis (in octaves), defined as $x = \log_2(F/F_0)$ where F_0 is the lower edge of the spectrum, i.e., 1, 0.5, or 0.25 kHz; and F is frequency. Φ is an arbitrarily chosen phase factor.

A moving ripple spectrum can be similarly characterized by its ripple frequency Ω (in cycles/octave), ripple phase Φ (in radians), and ripple velocity ω (in Hz).

$$S(x, t) = 1 + \Delta A \cdot \sin[2\pi(\omega \cdot t + \Omega \cdot x) + \Phi] \quad (2)$$

In these conventions, a positive value for ω corresponds to a ripple whose envelope travels toward low frequencies. Moving ripple stimuli lasted up to 1.7 s with similar rise-fall times. At the onset of the presentation, the ripple spectrum was initiated in a sine phase (defined as $\Phi = 0°$) relative to the low-frequency edge of the spectrum. The ripple began immediately moving to the left at the specified constant velocity, although the stimulus was only acoustically turned on 50 ms after the onset of motion. The overall level of a single-ripple stimulus was calculated from the level of a single frequency component at L_l dB SPL. Thus an L_l-level flat ripple is composed of 101 components, each at $L_l - 10 \log(101) \approx L_l - 20$ dB. The overall stimulus level was chosen on the basis of the threshold at the best frequency (BF); typically L_l was set at ~10 dB above threshold. High levels ($L_l > 65$ dB SPL) were avoided to ensure the linearity of our acoustic delivery system. For multiple-ripple stimuli, 100% modulation was defined by rescaling the complex profile so that its most negative peak just touches zero, i.e., analogous to setting $\Delta A = 1$ in the above equation.

This study concentrates on the responses to combinations of moving ripples. These were generated by first specifying the ripple frequency, initial phase, and velocity of each moving ripple in the combination. Then, the resulting compound waveform attributable to the superposition of these moving ripples was computed and used to shape the time-varying envelope of the complex stimulus. For example, Fig. 1 illustrates the spectral envelopes that result from adding different ripple combinations. The spectrogram in Fig. 1*A*, *left,* illustrates the envelope of the spectrum that results from adding a ripple with $\Omega = 0.8$ cycles/octave moving at 4 Hz, and starting at phase $\pi/6$, to another with $\Omega = 0.8$ cycles/octave moving at 8 Hz, and starting at π. At any instant, the spectral profile appears sinusoidal as a function of the logarithmic frequency axis with $\Omega = 0.8$ cycles/octave, as demonstrated by the cross section plotted above the spectrogram. The amplitude in time of any component in the spectrum is shown by the cross section at *right,* which is given by the sum of a 4-Hz and an 8-Hz sinusoid. Figure 1*B* displays a complex spectrogram resulting from the addition of five ripples at the velocities and random initial phases indicated in the figure legend.

Recordings

Action potentials from single units were recorded with the use of 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 on a computer. The computer also controlled stimulus delivery and created various raster displays of the responses. Each single-ripple stimulus combination was presented 40 times in a test, and multiple-ripple stimuli were usually presented 100–150 times.

A single unit was visually identified and isolated with the use of a windowing discriminator. Clusters are defined to be a group of two to five single units distinguished by spike amplitude. In each animal, electrode penetrations were made perpendicular to the cortical surface. Within a penetration, cells were typically isolated at depths of 350–600 μm, corresponding to cortical layers III and IV (Shamma et al. 1993).

Data analysis for tonal stimuli

For each cell, a frequency response curve was measured 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 as measured by counting the spikes evoked by the tone. The rate-level function at BF was measured at a range from 35 to 85 dB 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.

Data analysis for ripple stimuli

SINGLE STATIONARY RIPPLES. Each unit was initially tested with stationary single-ripple stimuli over the range of 0–2 cycles/octave in steps of 0.4 cycles/octave. At each ripple frequency, the amplitude and phase of the primary response component synchronized to the ripple frequency were determined (as described in detail in Shamma et al. 1995). The transfer function was then inverse Fourier transformed to compute the unit's $R\mathcal{T}$. The transfer function usually peaks around a characteristic ripple frequency, which will be referred to as Ω_o.

SINGLE MOVING RIPPLES. Single moving ripples were presented in one of two ways: *1*) over a range of velocities at a specified ripple frequency (usually at Ω_o) to measure the "temporal" transfer function, or *2*) over a range of ripple frequencies at a specific velocity (usually at ω_m, where the temporal transfer function is maximum) to measure the "ripple" transfer function. For either test, the strength of the phase-locked responses was assessed from period histograms with a time base of 16 or 32 bins constructed at each ω or Ω as described in detail in Kowalski et al. (1996). The amplitude and phase of the response component synchronized to each ω or Ω was then derived from the first coefficient of a 16- or 32-point fast Fourier transform of the histogram [$AC_1(\omega)$ or $AC_1(\Omega)$]. The amplitude of this component was then weighted by the total root-mean-square value of the response and used to construct the temporal or ripple transfer function of the unit [$T_{\Omega_o}(\omega)$ and $T_{\omega_m}(\Omega)$].

The temporal transfer function is therefore given by

$$T_{\Omega_o}(\omega) = AC_1(\omega) \cdot \frac{|AC_1(\omega)|}{\sqrt{\sum_{i=1}^{8} |AC_i(\omega)|^2}} \quad (3)$$

where $|AC_i(\omega)|$ is the magnitude of the ith Fourier component of the period histogram response. In general, $T_{\Omega_o}(\omega)$ can also be written as

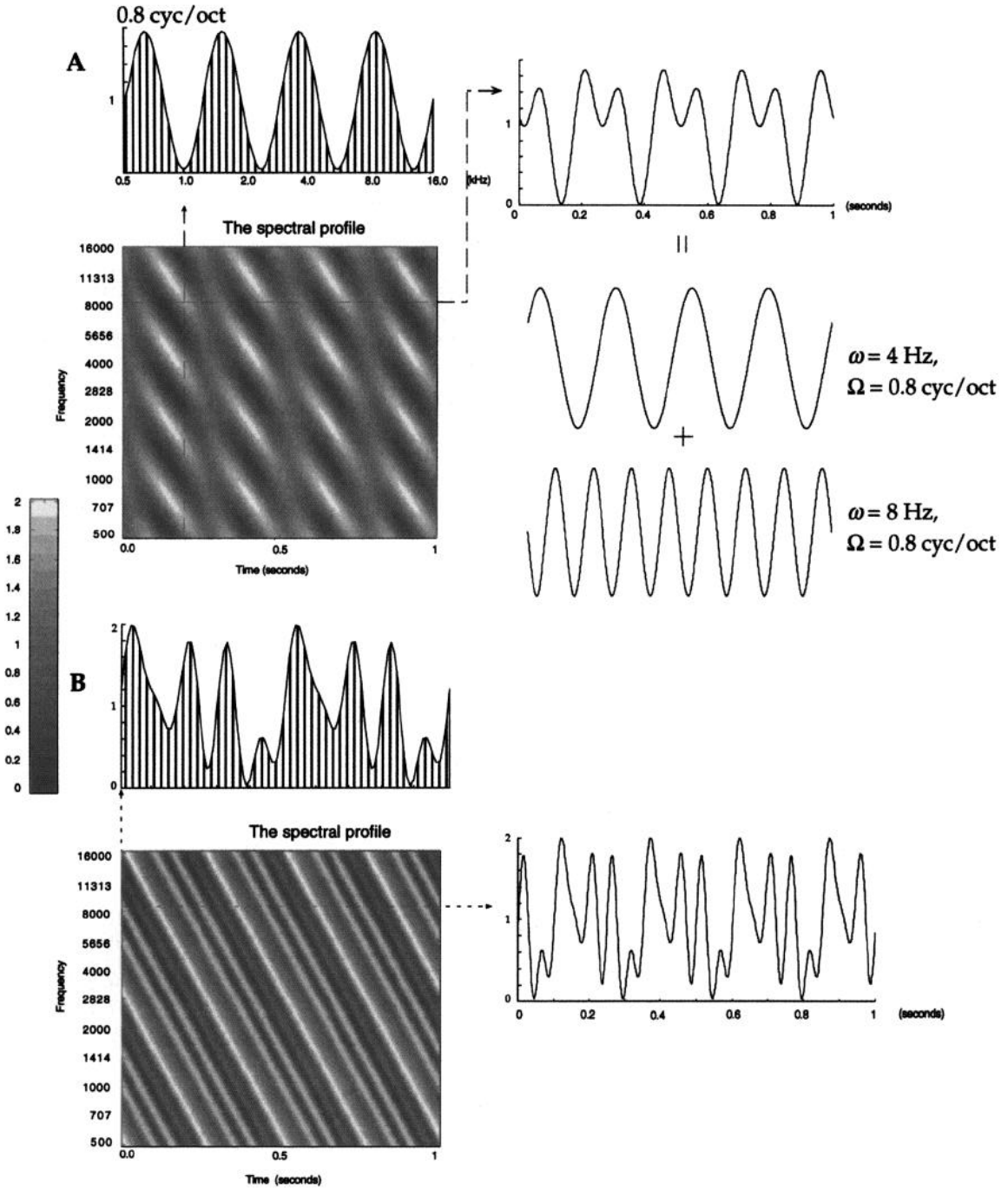

FIG. 1. Spectrograms of stimuli consisting of multiple moving ripple spectra. Each moving ripple is characterized by its ripple frequency (Ω), initial phase (Φ), and velocity (ω). *A*: spectrogram of a stimulus consisting of 2 moving ripples: Ω = 0.8 cycles/octave, ω = 4 Hz, $\Phi = \pi/6$ added to another ripple with Ω = 0.8 cycles/octave, ω = 8 Hz, and $\Phi = \pi$. The spectrogram of the resulting combination ripple is sinusoidal at every instant, with Ω = 0.8 cycles/octave as illustrated by the *top cross section.* The time course of each spectral component is demonstrated by the *right cross section,* which is the sum of 2 sinusoidal waveforms at 4 and 8 Hz. *B*: spectrogram of the addition of 5 ripples with random phases. The constituent ripples are: Ω = 0.4, 0.8, 1.2, 1.6, and 2.0 cycles/octave, ω = 4, 8, 12, 16, and 20 Hz, Φ = 0, 30, 41, −101, and 25°.

$$T_{\Omega_o}(\omega) = |T_{\Omega_o}(\omega)| e^{j\Phi_{\Omega_o}(\omega)} \quad (4)$$

where $j = \sqrt{-1}$. Figure 2 illustrates the magnitude $|T_{\Omega_o}(\omega)|$ and the unwrapped phase $\Phi_{\Omega_o}(\omega)$ of the transfer function $T_{\Omega_o}(\omega)$ (*top*). In almost all units recorded, the phase function could be fit well by a straight line $\hat{\Phi}_{\Omega_o}(\omega)$ (Kowalski et al. 1995). The slope of the line reflects the absolute time delay τ_d between stimulus and responses. Note that this estimate includes the additional delay attributable to the arbitrary choice of the starting time of the period histogram. In all cases shown in this paper, the period histograms are constructed from responses starting at t = 120 ms, and thus the absolute time delay can be computed from

$$\tau_d = \text{slope (radian/Hz)} - 0.12 \text{ s} \quad (5)$$

Another parameter of the linear fit of the phase function is its intercept along the ordinate, $\hat{\Phi}_{\Omega_o}(0)$, which is an additional con-

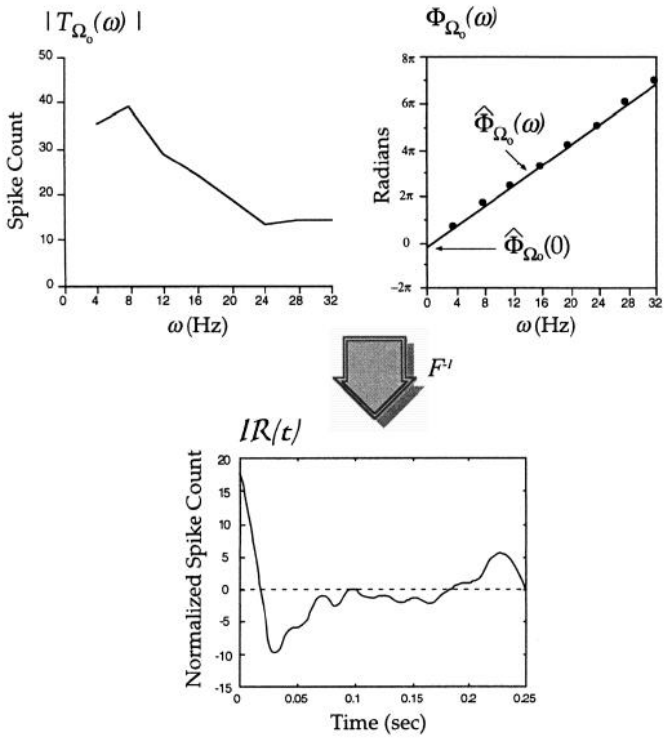

FIG. 2. Measuring the $\mathcal{IR}$ of a unit through its temporal transfer function $[T_{\Omega_o}(\omega)]$. Responses are measured over a range of temporal frequencies ω with a fixed ripple frequency Ω_o (the characteristic ripple of the unit) to determine the temporal transfer function $T_{\Omega_o}(\omega)$. *Top*: magnitude $|T_{\Omega_o}(\omega)|$ and unwrapped phase $\Phi_{\Omega_o}(\omega)$ of the transfer function. A straight line fit to the phase function, and its intercept $\hat{\Phi}_{\Omega_o}(0)$, are also shown. The phase function is then adjusted (see text for details) and the $T_{\Omega_o}(\omega)$ is inverse Fourier transformed to determine the impulse response function $\mathcal{IR}$ shown at the *bottom*.

stant phase shift in the period histogram relative to the stimulus ripple phase. For more details of this analysis, see Kowalski et al. (1996).

The ripple transfer function $T_{\omega_m}(\Omega)$ can be similarly written as

$$T_{\omega_m}(\Omega) = |T_{\omega_m}(\Omega)| e^{j\Phi_{\omega_m}(\Omega)} \quad (6)$$

where $j = \sqrt{-1}$. Figure 3 illustrates the magnitude $|T_{\omega_m}(\Omega)|$ and the unwrapped phase $\Phi_{\omega_m}(\Omega)$ of the transfer function $T_{\omega_m}(\Omega)$, respectively (*top*). A straight line fit to the phase function, $\Phi_{\omega_m}(\Omega)$, can be described by

$$\hat{\Phi}_{\omega_m}(\Omega) = x_m \cdot \Omega + \hat{\Phi}_{\omega_m}(0) \quad (7)$$

where x_m is the slope of the line and $\Phi_{\omega_m}(0)$ is its intercept. The parameter x_m reflects the location (in octaves) of the response field relative to the left edge of the ripple. The distance from the center of the response field envelope to the left edge of the spectrum is given by $k \cdot 2\pi/\Delta + x_m$, where Δ is the step size of the ripple frequencies tested and k is an integer ≥ 1 (Shamma et al. 1995). The intercept $\hat{\Phi}_{\omega_m}(0)$ is an additional constant phase shift in the period histogram relative to the stimulus ripple phase.

In most units, the transfer functions were measured only at one overall stimulus level that elicited a relatively strong response. Previous studies have determined that responses to stationary and traveling stimuli are not critically dependent on the base intensity (Kowalski et al. 1996; Shamma et al. 1995).

The ripple and temporal transfer functions were inverse Fourier transformed to obtain the corresponding $\mathcal{RF}$ and $\mathcal{IR}$ functions, shown respectively on the *bottom* of Figs. 2 and 3. In either case, the phase function was modified to remove the inappropriate constant phase shifts (Kowalski et al. 1996).

COMBINATIONS OF MOVING RIPPLES. Responses to stimuli composed of multiple moving ripples were recorded and compared with predictions made from the temporal and ripple transfer functions of the cell. Two types of moving ripple combination stimuli were presented. *1*) Relatively simple combinations of two to four ripples, all with the same ripple frequency but with different velocities: with four or fewer ripples, it is still possible to discern visually the contribution of each ripple to the spectral envelope. This stimulus allowed us to test linearity of the dynamic response, thus complementing the study of linearity in the spectral domain (Shamma and Versnel 1995). *2*) The second type of stimuli were more complex, comprising 5–15 moving ripples, with various combinations of ripple frequencies, velocities, and initial phases. For example, one temporally complex stimulus consists of up to 15 moving ripples with the same ripple frequency, but different velocities and random initial phases; another consists of up to 15 moving ripples with the same velocity, but different ripple frequencies and random initial phases. A third conceptually simple stimulus is a frequency-modulated tone type of spectrum, where the frequency and velocity of each ripple and its initial ripple phase are chosen so that the resulting composite spectrum at any instant consists of a single sharp peak sweeping downward at a constant velocity.

Measured responses from 100 repetitions of the stimulus were analyzed with the use of 16- or 32-bin period histograms with the period being the fundamental period of the stimulus. If the maximum spike count did not exceed 15 spikes in any of the bins, the response was considered weak and not considered for any further analysis.

Predicted responses of a unit were computed from its responses to single moving ripples, i.e., its temporal and ripple transfer functions, and their inverses the $\mathcal{IR}$ and $\mathcal{RF}$. For those stimuli in which only a single ripple frequency Ω was used (1st and 3rd categories above), only the temporal transfer function at that Ω was needed because the predicted curve could be directly extracted by superposition of the normalized period histograms measured at the appropriate velocities.

In general, however, predictions for all stimuli could be derived from the $\mathcal{RF}$ and $\mathcal{IR}$ of a unit as illustrated in Fig. 4. In Fig. 4*A*, the envelope of a three-ripple combination stimulus is depicted in the form of a spectrogram. The stimulus (and thus all responses) is periodic with a fundamental period of 250 ms. The $\mathcal{RF}$ of the cell is computed from the inverse Fourier transform of the phase-corrected ripple transfer function as discussed earlier, and is shown in Fig. 4*B* oriented (vertically) along the frequency (tonotopic) axis, with BF $\approx$ 3 kHz. Figure 4*C* illustrates the product of the $\mathcal{RF}$ with the stimulus profile as a function of time, which represents the response of the unit due to the $\mathcal{RF}$ alone. This (periodic) function is then modified by the dynamic response properties of the cell through a convolution with the $\mathcal{IR}$, shown in Fig. 4*D*. One fundamental period of the final predicted response of the unit is illustrated in Fig. 4*E* (bold line), superimposed (with an arbitrary scale) against the measured response of the cell to the stimulus (thin line). Note that, although the predicted curve fluctuates around zero (little or no sustained responses are usually observed), the measured response curve is always half-wave rectified (and sometimes also saturated).

To assess objectively the similarity between the two functions, a correlation coefficient is defined as

$$\rho = \frac{\Sigma_t r_{\text{meas}}(t) \cdot r_{\text{pred}}(t)}{\sqrt{\Sigma_t r^2_{\text{meas}}(t) \cdot \Sigma_t r^2_{\text{pred}}(t)}} \quad (8)$$

where $\tau_{\text{meas}}(t)$ is the measured spike count curve and $\tau_{\text{pred}}(t)$ is the predicted response curve. Because $\tau_{\text{meas}}(t)$ is half-wave rectified, the comparison would be more accurate if the correlation coefficient were computed with a half-wave-rectified $\tau_{\text{pred}}(t)$. In this case, uncorrelated responses have coefficients of $\sim$0.35 (rather than 0 for nonrectified functions). A histogram of the correlation coefficients from all units/clusters tested is compiled on the basis of the different categories of tests. Sometimes, several stimuli of the same type were presented to a cell, e.g., various combinations

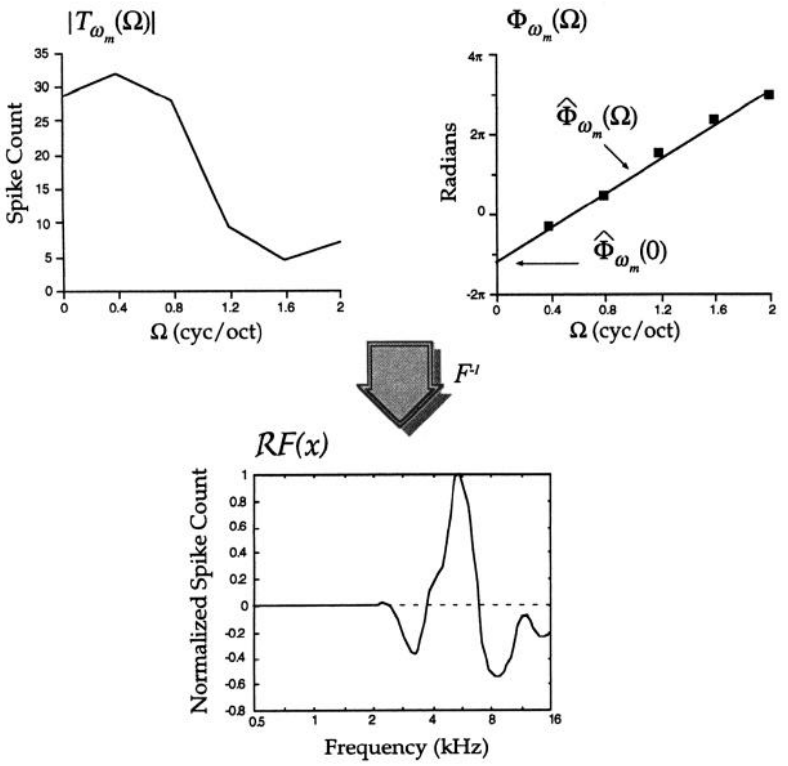

FIG. 3. Measuring the $\mathcal{RF}$ of a unit through its ripple transfer function $[T_{\omega_m}(\Omega)]$. Responses are measured over a range of ripple frequencies Ω with a fixed ripple velocity ω_m (the velocity of maximum response) to determine the ripple transfer function $T_{\omega_m}(\Omega)$. *Top*: magnitude $|T_{\omega_m}(\Omega)|$ and unwrapped phase $\Phi_{\omega_m}(\Omega)$ of the transfer function. A straight line fit to the phase function, and its intercept $\hat{\Phi}_{\omega_m}(0)$, are also shown. The phase function is then adjusted (see text for details) and $T_{\omega_m}(\Omega)$ is inverse Fourier transformed to determine the response field $\mathcal{RF}$ shown at the *bottom*.

of moving ripple pairs. In these cases, the correlation coefficient indicated is the average obtained from all such tests.

RESULTS

Two hundred ninety-three combination stimuli were presented to 51 single units and clusters in five animals (35 single units, 16 clusters). All examples shown in the figures are responses of single units, although those obtained from clusters were very similar in character. Correlation coefficients from the two populations are displayed separately in the histograms.

Responses to simple combinations of ripples: pairs, triplets and quadruplets

The response patterns and their predictions are illustrated in Fig. 5 for three different units stimulated by different two-ripple combinations. Spectrograms of the stimulus envelopes are shown at *left* (Fig. 5*A*). The $\mathcal{RF}$ and $\mathcal{IR}$ of each unit are

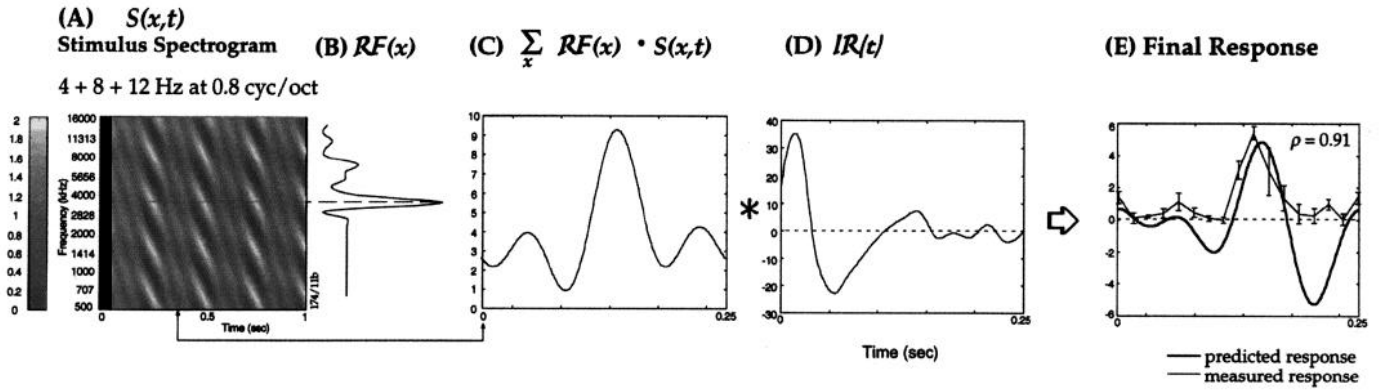

FIG. 4. Predicting the final response to multiple moving ripple stimuli. *A*: spectrogram of the stimulus $S(x,t)$, along with its ripple content (3 moving ripples in this case, all at $\Phi = 0$). Gray scale indicates relative amplitude of the spectrogram. *B*: $\mathcal{RF}$ of the cell (measured as described in Fig. 3) with best frequency (BF) = 3 kHz. The function is plotted sideways, i.e., aligned to the logarithmic frequency axis of the spectrogram (which also represents the tonotopic axis). *C*: product of the stimulus spectrogram and its $\mathcal{RF}$ generates a time function that is the response of the unit attributable to the $\mathcal{RF}$ alone. *D*: $\mathcal{IR}(t)$ of the cell (measured as described in Fig. 2). The $\mathcal{IR}$ is convolved with the function in *C* to produce the final response shown in *E*. *E*: final predicted response of the cell (thick solid line) superimposed on the measured spike count (thin solid line with error bars). The error bars (indicating mean ± SD) for the measured response curve and the correlation between measured and rectified predicted responses are also shown. Dashed line: 0 spike count. All abscissas measure time in s; *Y*-axes labeled on the left: normalized spike counts. *Y*-axes on the right side: actual spike count. Arrows: location of $t = 0$ of the periodic functions in *C*–*E*, relative to the corresponding period of the stimulus.

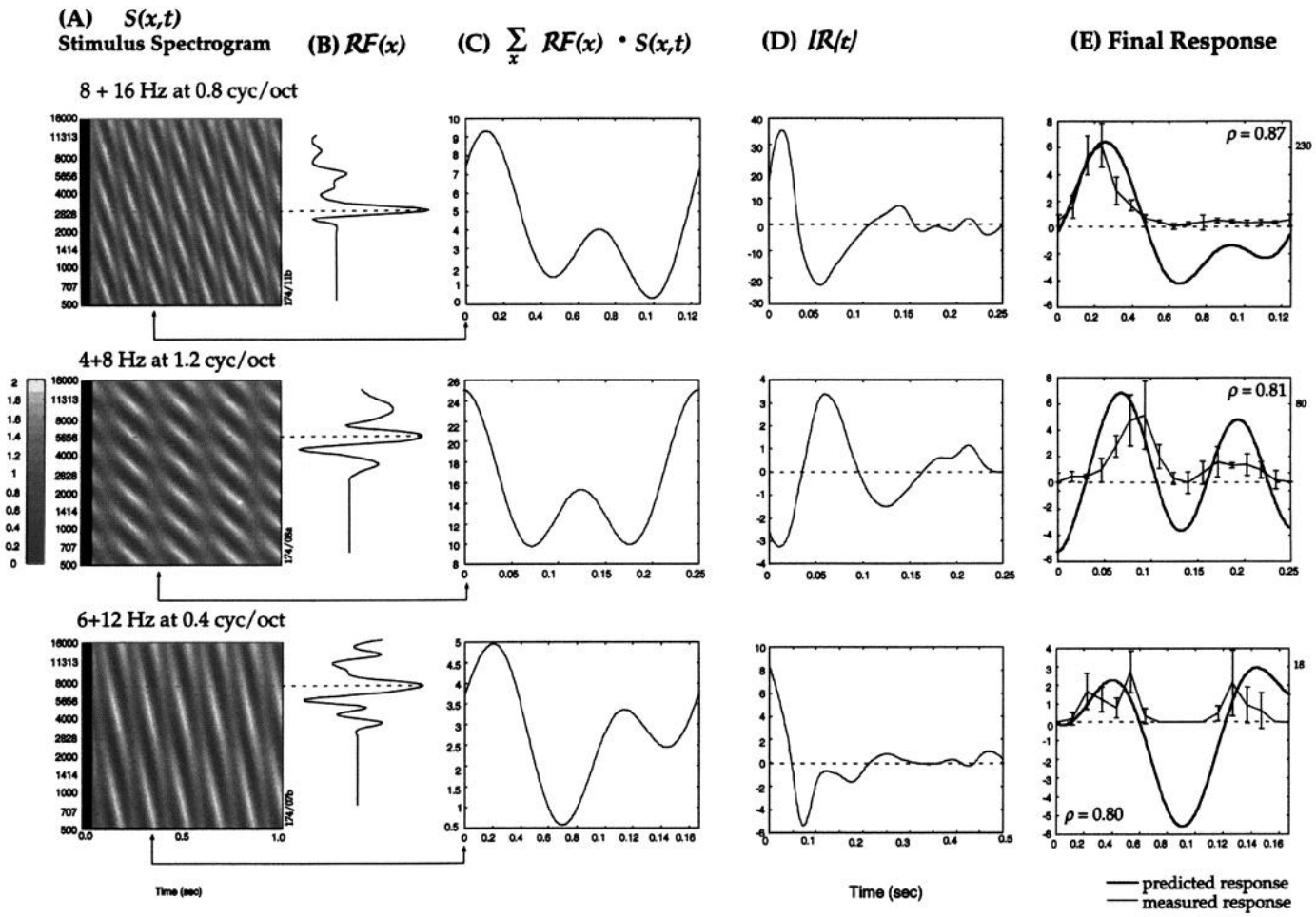

FIG. 5. Examples of responses to pairs of ripples from 3 cells. All details of the plots and computations are as in Fig. 4.

shown in *B* and *D*, respectively. Figure 5*C* illustrates the predicted response of each unit attributable to the $\mathcal{RF}$ alone, whereas the final predicted and measured responses are shown in *E*.

The ability of the $\mathcal{IR}$ to predict the dynamics of the responses is best seen by comparing *C* and *E* for each unit. For example, in the unit shown by the *middle row,* the responses attributable to the $\mathcal{RF}$ alone (*C*) become almost completely inverted in *E* (bold curve). This inversion is due to the unusual "inverted" form of the $\mathcal{IR}$, which is described in detail in the companion paper (Kowalski et al. 1996). More typical $\mathcal{IR}$s as in the *top* and *bottom* units produce similar, although not as dramatic, transformations of the waveforms between *C* and *E,* such as additional absolute delays (reflecting the effects of τ_d), and changes in the relative heights of the response peaks, most notably in the *bottom* unit, where, in the final response (*E*), the two peaks become comparable in size.

Figure 6 shows responses from the same unit as in Fig. 5 (*top*) to three stimuli with increasing numbers of moving ripples. The responses generally exhibit the same features as before, especially the waveform transformations attributable to the $\mathcal{IR}$ (*C* vs. *E*). Furthermore, the predicted responses (bold lines in *E*) match reasonably well the outlines and some of the details of the measured responses (thin lines) ($\rho > 0.78$). For instance, time of occurrence of the largest peak in the response varies depending on the stimulus in a similar manner for both predicted and measured responses. Another example of such a matching covariation of the responses for different stimuli is shown in Fig. 7. In addition, the $\mathcal{IR}$ of this unit (*D*) is such that it induces a significant transformation of the responses similar to that seen earlier for the *middle* unit in Fig. 5. Note, however, that in both examples of Figs. 5 and 6, predicted responses may be delayed or advanced relative to the measured curve. Such shifts are discussed later in the context of possible measurement errors.

In general, the most prominent disparity between predicted and measured response curves in all cases is due to the half-wave rectification of the spike rates (Kowalski et al. 1996). The effects of this instantaneous nonlinearity can be readily understood by comparing the responses to the stimulus and its inverted envelope shown in Fig. 8. In the *top row,* the final response has a single large peak that occurs at 170 ms. In the *bottom row,* the stimulus envelope has been inverted. The response also becomes inverted, thus revealing the previously half-wave rectified fluctuations. Given the two responses, one can construct a "linearized" or nonrectified version (dashed curve in *E, bottom*) that matches better the linear predictions (——).

A histogram of the correlation coefficients generated by the responses to stimuli with two, three, or four ripples and their predictions is displayed in Fig. 9. Most responses to these ripple stimuli were reasonably predictable (84% with $\rho > 0.6$).

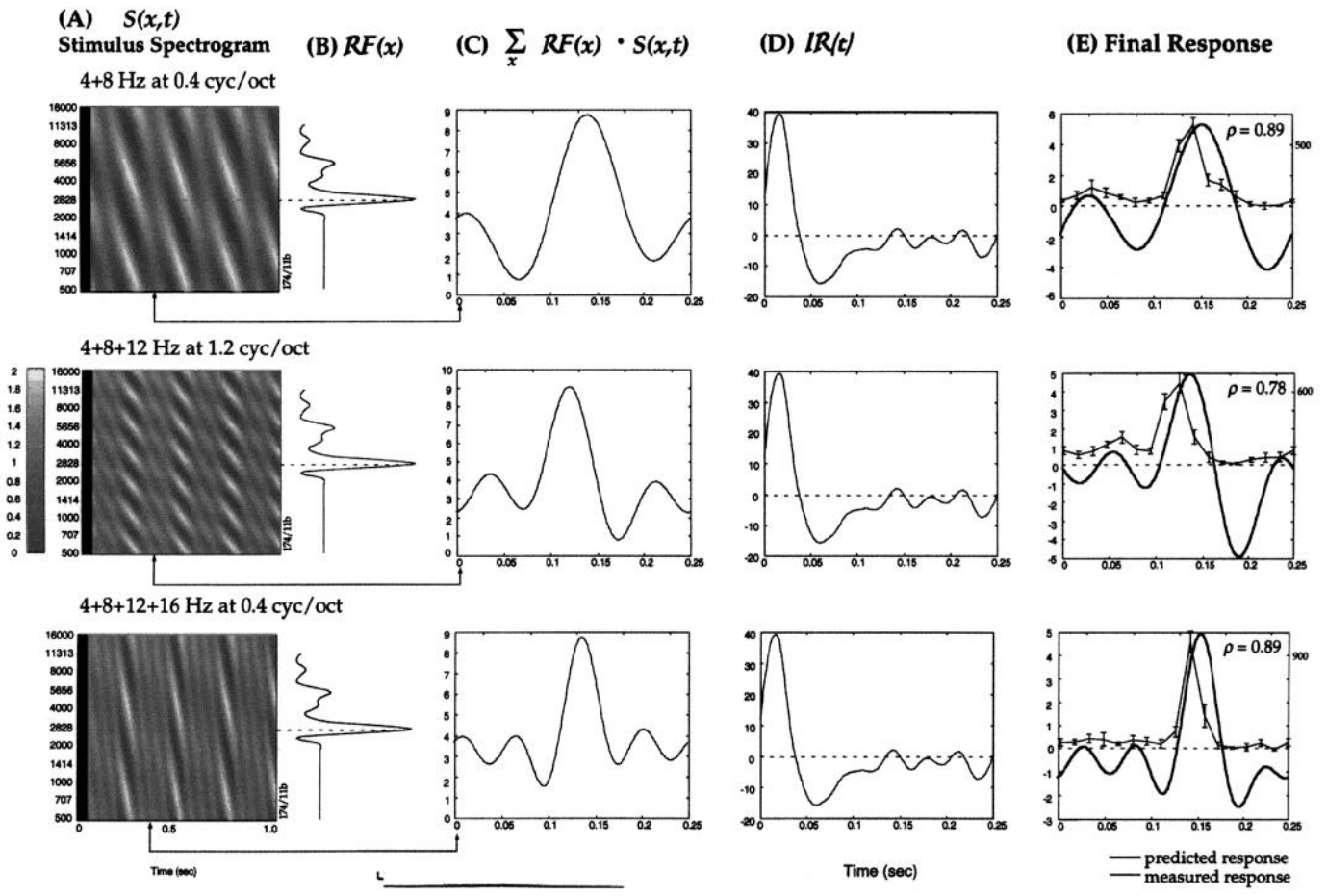

FIG. 6. Examples of responses of a unit to 3 combination ripple stimuli: *top row,* 2 ripples; *middle row,* 3 ripples; *bottom row,* 4 ripples. All details of the plots and computations are as in Fig. 4.

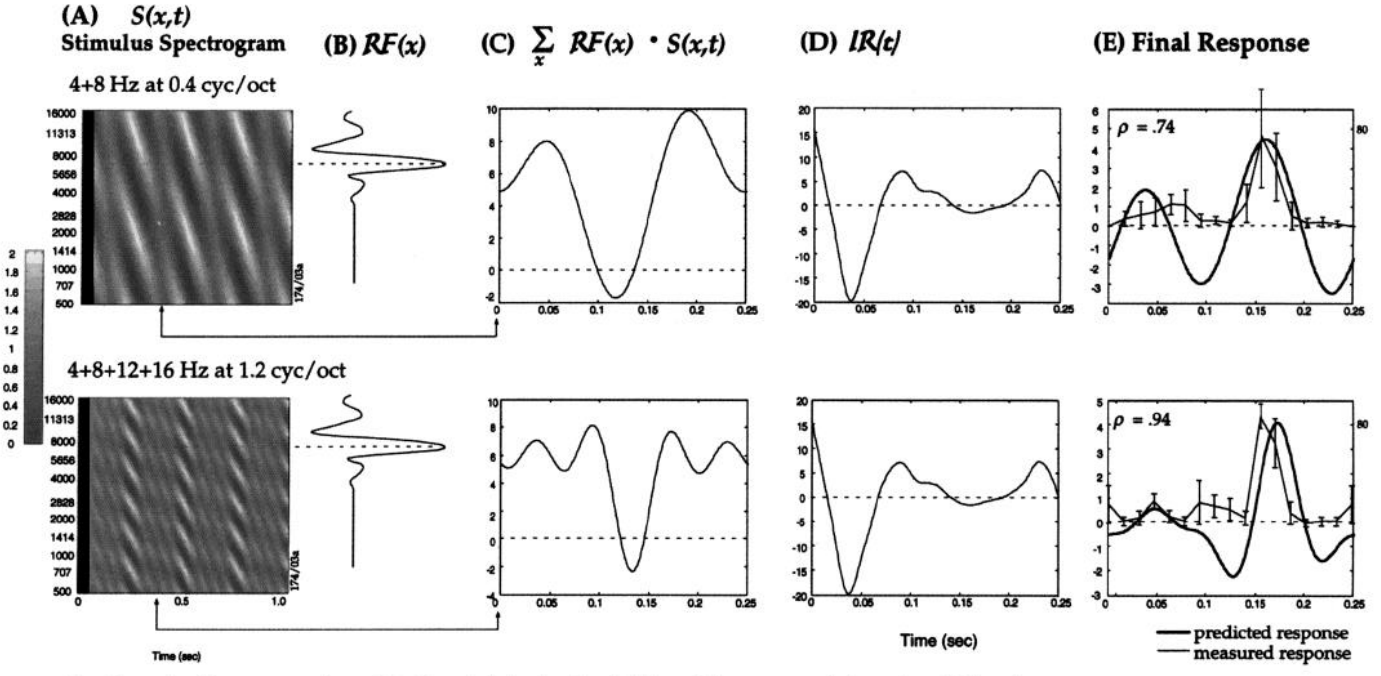

FIG. 7. Example of responses of a unit to 2- and 4-ripple stimuli. The unit has a strongly inverting $\mathcal{HR}$ function.

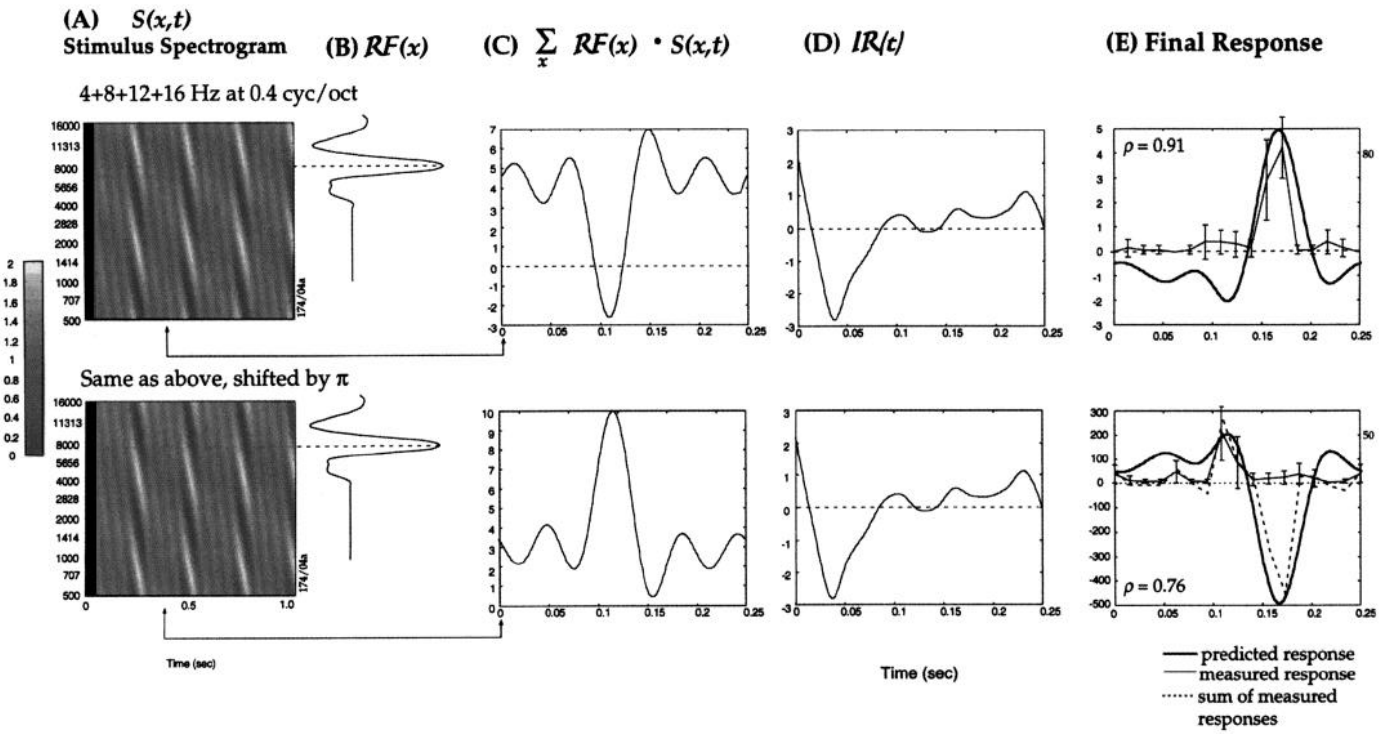

FIG. 8. Uncovering the half-wave rectified portion of the responses. The 2 stimuli illustrated are inverted relative to each other, and thus their responses are also inverted. The measured and predicted responses are indicated as described in Fig. 4 above. The measured responses are also used to construct the "nonrectified" version of the response, depicted by the dashed curve in *E, bottom.* All other details of the plots and computations are as in Fig. 4.

Responses to complex combinations of ripples

The results shown in Figs. 10 and 11 demonstrate further the predictive power of the $\mathcal{RF}$ and $\mathcal{IR}$ and thus the extent of the linearity of the responses. Figure 10 shows examples of responses and predictions to complex combinations of moving ripples for one single unit. As many as 15 ripples were added together to create these stimuli, resulting in complex envelope profiles as seen in the spectrograms (Fig. 10*A*). The *top* stimulus has a complex temporal structure at any given frequency, but is spectrally simple, whereas the *bottom* stimulus has the opposite character. The stimulus in the *middle* has a frequency-modulated-like character, mimicking the spectrum of a single traveling tone. Predictions in all cases give significant correlation coefficients (0.88, 0.83, 0.89).

Figure 11 illustrates similar findings from additional complex stimuli or different units. Note that for all three examples, the impulse response (Fig. 11*D*) plays a key role in that the final responses cannot be readily predicted from the $\mathcal{RF}$ alone. Instead, the waveforms in Fig. 11*C* are almost inverted by the $\mathcal{IR}$ in the top two examples; additionally, in the *middle* example, the fast temporal modulations in the input and in *C* are heavily filtered out by the $\mathcal{IR}$ to produce finally a smoother response (*E*).

The histogram in Fig. 12 sums up the correlation coefficients found between predicted and measured responses for all units tested with complex ripple combinations. As with simple ripple combinations, the majority of correlation coefficients are significantly large (89% with $\rho > 0.6$).

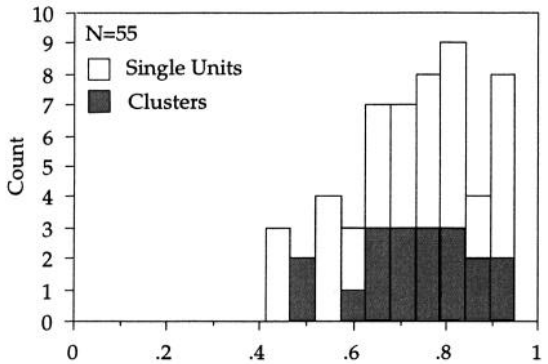

FIG. 9. Histogram of the correlation coefficients between measured and predicted responses to stimuli with 2, 3, or 4 moving ripples.

DISCUSSION

Summary of responses to moving ripple combinations

The results presented support the hypothesis that AI units responses to arbitrary dynamic spectra are reasonably predictable once the response to single moving ripples is known, or specifically, once the $\mathcal{RF}$ and $\mathcal{IR}$ are known. This was demonstrated by validating the superposition principle, that is, the responses to a combination of moving ripples compare well with those predicted from a linear sum of the responses to the individual constituent ripples. This was found to hold for spectra composed of up to 15 ripples, with various ripple frequencies, velocities, and initial phases, and despite a wide range of sources for error. For instance, some error is introduced by the assumption of separability of the spectral and temporal dimensions of the transfer function, as discussed in Kowalski et al. (1996). Errors are also inevitable in com-

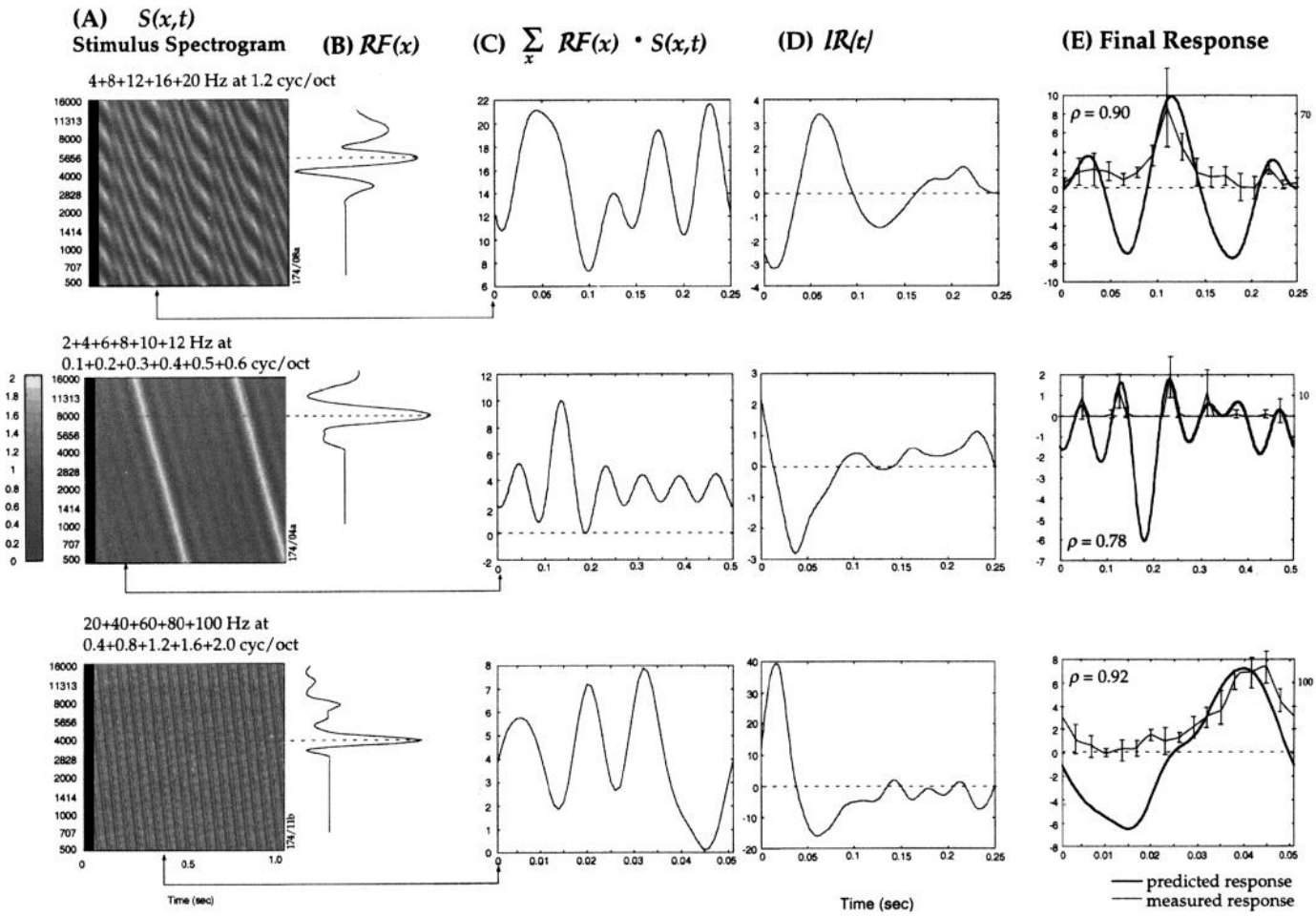

FIG. 10. Responses of 3 different cells to 3 complex combinations. In all cases, the $\mathcal{IR}$ plays a key role in shaping the final responses of the cells. The stimuli used are: *top row,* temporal noise composed of 5 ripples; *middle row,* a simulated frequency-modulated stimulus effectively traveling at 20 octaves/s; *bottom row,* a simulated frequency-modulated stimulus effectively traveling at 50 octaves/s. All other details of the plots and computations are as in Fig. 4.

puting the $\mathcal{RF}$ and $\mathcal{IR}$, where the phase functions have to be fitted and adjusted to subtract the inappropriate terms (as discussed in METHODS and in Kowalski et al. 1996). Another apparent error is the shift (advance or delay) of the predicted response curve relative to the measured one, as in Figs. 5 and 6. These are most likely due to inaccuracies in the measurements of the slope of the temporal phase function $\Phi_{\Omega_o}(\omega)$ that affect the τ_d, the absolute time delay of the $\mathcal{IR}$. Finally, measurements of the transfer functions and of the multiple-ripple stimuli are performed sequentially over a relatively long period of time (from 30 to 60 min), during which the state of the animal is likely to change somewhat.

Effects of response nonlinearities

AI units exhibit response nonlinearities attributable to threshold and saturation that cause their firing rates to be half-wave rectified, clipped, or exponentially distorted. The first two are well-recognized properties of nerve cell firings and their effects are immediately visible on comparing the (linearly) predicted response waveforms (which can go negative) with the actual responses shown in most examples in this paper. The third effect is more subtle; it is well illustrated by the responses in Fig. 6, where the response peaks appear to be sharper (or to have steeper skirts), as if the predicted waveform has been exponentially enhanced. This distortion is common in many responses shown here and in Kowalski et al. (1996).

However, these largely instantaneous nonlinearities appear to act on the already generated linear response pattern. Therefore their effects are relatively transparent and the underlying linearly predicted waveform is immediately accessible as demonstrated in all examples in the paper. Furthermore, the information content in the distorted response waveform remains intact and thus the linear version of the response can be recovered easily in a manner similar to that discussed in Shamma et al. (1995) for responses to stationary spectra. These conclusions are consistent with findings by Nelken et al. (1994) that responses to complex sounds (e.g., 9-tone complexes) are predictable on the basis of the response areas measured with two-tone stimuli.

Functional significance of AI temporal response properties

AI units respond to changes in the spectral envelope in a substantially linear and temporally selective manner. They are usually tuned around a specific rate between 2 and 16 Hz, with an approximate bandwidth of 3 octaves. These response parameters are summarized for all units isolated in

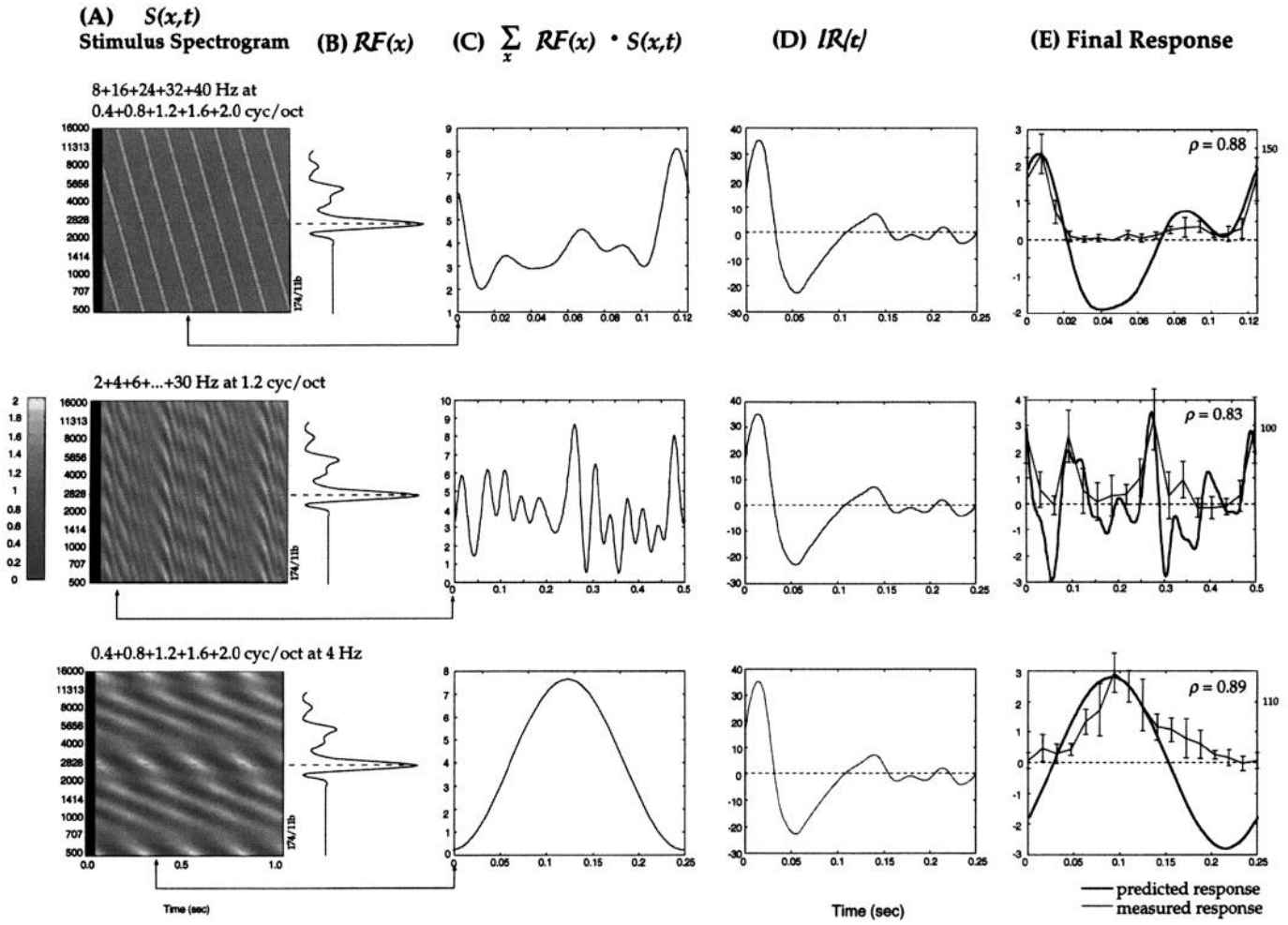

FIG. 11. Examples of responses to complex combinations of ripples. Three complex combinations are presented to the same unit: *top row,* simulated frequency-modulated stimulus composed of 5 ripples to produce effectively a single peak with velocity 20 octaves/s; *middle row,* temporal noise stimulus generated with 15 random-phase moving ripples all at $\Omega = 1.2$ cycles/octave; *bottom row,* ripple noise stimulus generated by adding 5 ripples with different Ωs and random phases, all moving at the same velocity. All other details of the plots and computations are as in Fig. 4.

Fig. 6 in the companion paper (Kowalski et al. 1996). The functional implications of the histogram distribution, however, are ambiguous. On the one hand, AI units can be said to be all tuned around an average rate of 8 Hz; the scatter around this value in the histogram is then the usual noise one expects in a physiological system. On the other hand, one may interpret the histogram as a broad distribution of cells with transfer functions tuned to different temporal rates, all with approximately similar bandwidths.

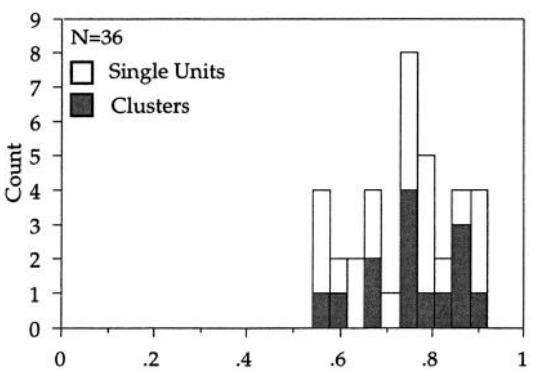

FIG. 12. Histogram of the correlation coefficients between measured and predicted responses to complex ripple stimuli with ≥5 moving ripples.

In the first view, the 8-Hz tuning may be seen as a general physiological limitation of cortical cell dynamics, or as an epiphenomenon related to the rhythms induced by corticothalamic loops (Eggermont 1992). It may also be that temporal tuning per se is not important, but rather that the impulse responses ($\mathcal{IR}$) act functionally as a temporal derivative that abolishes the sustained responses to stationary spectra and preserves only responses to dynamic spectra.

The second interpretation of the histogram implies that AI units have impulse response functions with a range of dilations analogous to the range of different bandwidths exhibited by the $\mathcal{RF}$s. Specifically, it is assumed in this hypothesis that for any given $\mathcal{RF}$, there are different units with a range of $\mathcal{IR}$s, each encoding the local dynamics of the spectrum at a different time scale, i.e., there are units exclusively sensitive to slow modulations in the spectrum, and others tuned to moderate or fast changes. This temporal decomposition is analogous to the multiscale representation of the shape of the spectrum produced by the RFs. Such an analysis may underlie many important perceptual invariances such as the ability to recognize speech and melodies despite large

changes in rate of delivery (Julesz and Hirsh 1972), or to perceive continuous music and speech through gaps, noise, and other short-duration interruptions in the sound stream. Furthermore, the segregation into different time scales such as fast and slow corresponds to the intuitive classification of many natural sounds and music into transient and sustained categories, or into stop consonants and continuents in speech.

Finally, an overall view of the AI representation of a dynamic spectrum can be summarized as follows. First, AI units with a wide range of $\mathcal{RF}$s generate a multiscale representation of the spectrum. Next, the dynamic responses of each unit are effectively "differentiated" in time by the $\mathcal{IR}$, possibly with different degrees of temporal resolution. This combined spectrotemporal decomposition is remarkably similar in spirit to the common practice in engineering systems (such as speech recognition systems) (Rabiner and Schafer 1978) in which speech spectra are represented in terms of cepstral coefficients and their temporal derivatives (the differential cepstral coefficients). As discussed in more detail in (Wang and Shamma 1995), the major conceptual difference between the two schemes is the multiscale nature of the AI representation.

We thank P. Gopalaswamy for help with the acquisition and analysis system, and A. Owens, H. Versnel, and T. Denison for assistance with animal preparation and data recordings. We also thank P. Cariani for insightful discussions, and two anonymous referees for helpful comments.

This work was supported by grants from the Office of Naval Research and the Air Force Office of Scientific Research, research grant NIDCD T32 DC-0046-01 from the National Institute on Deafness and Other Communication Disorders, and the National Science Foundation (NSFD CD 8803012).

Address for reprint requests: D. Depireux, Institute for Systems Research, University of Maryland, College Park, MD 20742-3311.

Received 21 December 1995; accepted in final form 18 July 1995.

REFERENCES

EGGERMONT, J. J. Stimulus induced and spontaneous rhythmic firing of single units in cat primary auditory cortex. *Hear. Res.* 61: 1–11, 1992.

EGGERMONT, J. J. Temporal modulation transfer functions for AM and FM stimuli in cat auditory cortex. Effect of carrier type, modulating waveform and intensity. *Hear. Res.* 74: 51–66, 1994.

EVANS, E. F. Single-unit studies of mammalian cochlear nerve. In: *Auditory investigations: The Scientific and Technological Basis,* edited by H. A. Beagley. Oxford, UK: Clarendon, 1979, p. 324–367.

JULESZ, B. AND HIRSH, I. J. Visual and auditory perception—an essay of comparison. In: *Human Communication: A Unified View,* edited by E. E. David, Jr. and P. B. Denes. New York: McGraw-Hill, 1972, p. 283–340.

KOWALSKI, N., DEPIREUX, D. A., AND SHAMMA, S. A. Analysis of dynamic spectra in ferret primary auditory cortex. I. Characteristics of single-unit responses to moving ripple spectra. *J Neurophysiol.* 76: 3503–3523, 1996.

NELKEN, I., PRUT, Y., VAADIA, E., AND ABELES M. Population responses to multifrequency sounds in the cat auditory cortex: One- and two-parameter families of sounds. *Hear. Res.* 72: 206–222 1994.

RABINER, L. R. AND SCHAFER, R. W. *Digital Processing of Speech Signals.* Englewood Cliffs, NJ: Prentice-Hall, 1978.

DE RIBAUPIERRE, F., GOLDSTEIN, M. H., AND YENI-KOMSHIAN, G. Cortical coding of repetitive acoustic pulses. *Brain. Res.* 48: 205–225, 1972.

SHAMMA, S. A., FLESHMAN, J. W., WISER, P. R., AND VERSNEL, H. Organization of response areas in ferret primary auditory cortex. *J. Neurophysiol.* 69: 367–383, 1993.

SHAMMA, S. A. AND VERSNEL, H. Ripple analysis in ferret primary auditory cortex. II. Prediction of unit responses to arbitrary spectral profiles. *Auditory Neurosci.* 1: 255–270, 1995.

SHAMMA, S. A., VERSNEL, H., AND KOWALSKI, N. Ripple analysis in ferret primary auditory cortex. I. Response characteristics of single units to sinusoidally rippled spectra. *Auditory Neurosci.* 1: 233–254, 1995.

WANG, K. AND SHAMMA, S. A. Spectral shape analysis in the central auditory system. *IEEE Trans. Speech Audio Proc.* 3: 382–395, 1995.