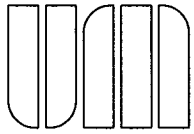
TECHNICAL RESEARCH REPORT

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

by N. Kowalski, D.A. Depireux and S.A. Shamma

T.R. 95-109

Sponsored by
the National Science Foundation
Engineering Research Center Program,
the University of Maryland,
Harvard University,
and Industry



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 & Electrical Engineering Department
University of Maryland
College Park, MD 20742-3311, USA
(301) 405-6842

1 Summary and Conclusions

1. Responses of single-units and unit clusters were recorded in the ferret primary auditory cortex (AI) using broadband complex dynamic spectra. Previous work (Kowalski et al 1995) 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. An arbitrary complex dynamic spectral profile can be conceptually thought of as being composed of a weighted sum of moving ripple spectra. Such a decompositon can be computed from a two-dimensional spectro-temporal 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.

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 to 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 suggest tha AI cells may encode the profile of a dynamic spectrum by performing a multiscale spectro-temporal decomposition of the dynamic spectral profile in a largely linear manner.

2 Introduction

Acoustic stimuli with broadband dynamic spectra evoke strong and relatively sustained responses in neurons of the primary auditory cortex (AI) (de Ribaupierre et al. 1972, Eggermont 1994, Kowalski et al. 1995). The response patterns reflect details of both the spectral shape and its changes in time. To characterize these neurons or units, elementary broad-band 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. 1995). A typical and most prominent feature of the responses is the synchronized component which 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 hence can be linearly weighted and summed to predict the responses to arbitrary combinations of stationary ripples (Shamma and Versnel 1995). This result also applies to moving ripples since the $\mathcal{RF}$ shape (to within a scaling factor) is independent of the ripple velocities (Kowalski et al. 1995).

However, the relevance of the $\mathcal{IR}$ in characterizing AI unit responses is uncertain since 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 is operable 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 to 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.

3 METHODS

3.1 Surgery and animal preparation

Ferrets (*Mustela putorius*) were anesthetized with sodium pentobarbital (40 mg/kg) and maintained in an areflexic state throughout the experiment by a continuous intravenous infusion of pentobarbital and lactated Ringer solution, mixed with dextrose to compensate for their high metabolic rate. The ectosylvian gyrus, which includes the primary auditory cortex, 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). The protocols for anesthesia and animal care were approved by the University's institutional animal care and use committee.

3.2 Acoustic stimuli

Various auditory stimuli were used: pure tones (single tone bursts, 200 ms duration, 10 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. 1995). 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 using a 1/8-in. Brüel & 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 broad-band 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:

$$S(x) = 1 + \Delta A \cdot \sin(2\pi \cdot \Omega \cdot x + \Phi), \tag{1}$$

where ΔA is 0.9 or 1, x is the logarithmic frequency axis (in octaves) defined as: $x = log_2(\frac{F}{F_o})$ where F_0 is the lower edge of the spectrum, i.e. 1 kHz, 0.5 kHz or 0.25 kHz, and F is frequency. Φ is an arbitrarily chosen phase factor. Note that when ΔA is zero, the resulting stimulus is a flat spectrum.

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 Hertz).

$$S(x,t) = 1 + \Delta A \cdot \sin(2\pi(\omega t + \Omega x) + \Phi) \ . \tag{2}$$

In these conventions, a positive value for ω corresponds to a ripple whose envelope travels towards low frequencies. Moving ripple stimuli lasted up to 1.7 seconds with similar rise/fall times. At the onset of the presentation, the ripple spectrum was initiated in a sine phase (defined as $\Phi = 0^o$) 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_1 dB SPL. Thus, an L_1 level flat ripple is composed of 101 components, each at $L_1 - 10\log(101) \approx L_1 - 20$ dB. The overall stimulus level was chosen on the basis of the threshold at BF, typically L_1 was set about 10 dB above threshold. High levels ($L_1 > 65$ dB SPL) were avoided to ensure the linearity of our acoustic delivery system. 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 concentrated 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 due to the superposition of these moving ripples was computed and used to shape the time-varying envelope of the complex stimulus. For example, Figure 1 illustrates the spectral envelopes that result from adding different ripple combinations. The spectrogram to the left in Fig.1A 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 $\frac{\pi}{6}$, to another with $\Omega = 0.8$ cycles/octave, velocity of 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 to the right, which is given by the sum of a 4 and 8 Hz sinusoids. Fig.1B displays a complex spectrogram resulting from the addition of 5 ripples at the velocities and random initial phases indicated in the figure legend.

3.3 Recordings

Action potentials from single units were recorded using glass-insulated tungsten microelectrodes with 5-6 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 150 times.

A single unit was visually identified and isolated using a windowing discriminator. Clusters are defined to be a group of 2-5 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).

3.4 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 best frequency (BF) was determined from this response curve as the frequency which evoked the best response as measured by counting the spikes evoked by the tone. Thus, BF is the frequency that elicits a response at the lowest possible threshold. The rate-level function at BF was measured at a range from 35 to 85 dB SPL in order to determine the cell's response threshold and the non-monotonicity. The criteria were 10% of maximum response and a decrease of 25% with increase of intensity, respectively.

3.5 Data analysis for ripple stimuli

3.5.1 Single stationary ripples

Each unit was initially tested with *stationary* single ripple stimuli over the range 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 $\mathcal{RF}$. The transfer function usually peaks around a characteristic ripple frequency, which will be referred to as Ω_o.

3.5.2 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. 1995). 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 FFT of the histogram ($AC_1(\omega)$ or $AC_1(\Omega)$).

The amplitude of this component was then weighted by the total *rms* 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:

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

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

$$T_{\Omega_o}(\omega) = |T_{\Omega_o}(\omega)| e^{j\Phi_{\Omega_o}(\omega)} , \quad \text{where } j = \sqrt{-1} . \tag{4}$$

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 panels). 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 is contaminated by the additional delay due 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 hence the absolute time delay can be computed from:

$$\tau_d = \text{slope (radian/Hz)} - 0.12 \text{ seconds.} \tag{5}$$

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 constant phase shift in the period histogram relative to the stimulus ripple phase. For more details of this analysis, see Kowalski et al. (1995).

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

$$T_{\omega_m}(\Omega) = |T_{\omega_m}(\Omega)| e^{j\Phi_{\omega_m}(\Omega)} , \quad \text{where } j = \sqrt{-1}. \tag{6}$$

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 panels). A straight line fit to the phase function, $\hat{\Phi}_{\omega_m}(\Omega)$ can be described by:

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

where x_m is the slope of the line, and $\hat{\Phi}_{\omega_m}(0)$ 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 RF envelope to the left edge of the spectrum is given by $k \cdot \frac{2\pi}{\Delta} + x_m$, 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 which elicited a relatively strong response. Previous studies have determined that responses to stationary and traveling stimuli are not critically dependent on the base intensity (Shamma et al. 1995; Kowalski 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 in the bottom panels of Figs.2 and 3. In either case, the phase function was modified to remove the inappropriate constant phase-shifts (Kowalski et al. 1995).

3.5.3 Combinations of moving ripples

Responses to stimuli composed of multiple moving ripples were recorded and compared to predictions made from the temporal and ripple transfer functions of the cell. The ripple-combination stimuli fell into one of the following categories: (1) 2-4 moving ripple combinations, all with the same ripple frequency, but with different velocities; (2) Simulated FM-tone spectrum composed of 5 moving ripples with ripple frequencies, initial ripple phases, and velocities chosen such that the composite spectrum resembles a single sharp peak sweeping downwards at a constant velocity; (3) Simulated temporal noise composed of 5-15 moving ripples with the same ripple frequency, but different velocities and random initial phases; (4) Simulated ripple noise composed of 5 moving ripples with the same velocity, but different ripple frequencies and random initial phases.

Measured responses from 100 repetitions of the stimulus were analyzed using 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 (categories 1 and 3 above), only the temporal transfer function at that Ω was needed since 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 Figure 4. In Fig. 4A, the envelope of a 3 ripple-combination stimulus is depicted in the form of a spectrogram. The stimulus (and hence 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.4B oriented (vertically) along the frequency (tonotopic) axis, with $BF \approx 3$ kHz. Fig.4C 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.4D. One fundamental period of the final predicted response of the unit is illustrated in Fig.4E (solid line), superimposed (with an arbitrary scale) against the measured response of the cell to the stimulus (dashed line). Note that, while the predicted curve fluctuates equally 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 :

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

where $r_{meas}(t)$ is the measured spike count curve and $r_{pred}(t)$ is the predicted response curve. Since $r_{meas}(t)$ is half-wave rectified, the comparison would be more accurate if the correlation coefficient were computed with a half-wave rectified $r_{pred}(t)$. In this case, uncorrelated responses have coefficients of about 0.35 (rather than zero for non-rectified functions). A histogram of the correlation coefficients from all units/clusters tested is compiled based on the different

categories of tests. Sometimes, several stimuli of the same type were presented to a cell, e.g., various combinations of moving ripple pairs. In these cases, the correlation coefficient indicated is the average obtained from all such tests.

4 RESULTS

293 combination stimuli were presented to 51 single units and clusters in 5 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.

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

The response patterns and their predictions are illustrated in Figure 5 for three different units stimulated by different two-ripple combinations. Spectrograms of the stimulus envelopes are shown in column A. The $\mathcal{RF}$ and $\mathcal{IR}$ of each unit are shown in columns B and D, respectively. Column C illustrates the predicted response of each unit due to the $\mathcal{RF}$ alone, while the final predicted and measured responses are shown in column E.

The ability of the $\mathcal{IR}$ to capture and predict the dynamics of the responses is best seen by comparing panels C and E for each unit. For example, in the middle unit, the responses due to the $\mathcal{RF}$ alone (panel C) become almost completely inverted in panel E (dashed 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. 1995). More typical $\mathcal{IR}$'s as in the top and bottom units produce similar, though not as dramatic, transformations of the waveforms between panels 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 (panel 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 due to the $\mathcal{IR}$ (panels C vs E). Furthermore, the predicted responses (solid lines in column E) match reasonably well the outlines and some of the details of the measured responses ($\rho > 0.78$). 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 co-variation of the responses for different stimuli is shown in Figure 7A. In addition, the $\mathcal{IR}$ of this unit (column 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 Fig.5 and 6, predicted responses may be delayed or advanced relative to the measured curve. Such shifts are most likely due to errors in the measurements of the slope of the temporal phase function $\Phi_{\Omega_o}(\omega)$ which affect the τ_d - the absolute time-delay of the $\mathcal{IR}$.

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. 1995). The

effects of this instantaneous nonlinearity can be readily understood by comparing the responses to the stimulus and its inverted envelope shown in Fig.7B. 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 non-rectified version (dotted curve in the bottom panel E) that matches better the linear predictions (solid lines).

A histogram of the correlation coefficients generated by the responses to stimuli with two, three or four ripples and their predictions are displayed in Figure 8. Most responses to these ripple stimuli were reasonably predictable (84% with $\rho > 0.6$).

4.2 Responses to complex combinations of ripples

The results shown in Figures 9 and 10 demonstrate further the predictive power of the $\mathcal{RF}$ and $\mathcal{IR}$ and hence the extent of the linearity of the responses. Figure 9 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 (column A). Stimulus at the top was an FM-like stimulus, mimicking the spectrum of a single traveling tone. The middle stimulus is temporally noise-like, but spectrally simple, whereas the last stimulus has the opposite character. Predictions in all cases give significant correlation coefficients (0.88, 0.83, 0.89).

Figure 10 illustrates similar findings from additional complex stimuli or different units. Note that for all three examples, the impulse response (Column D) plays a key role in that the final responses cannot be readily predicted from the $\mathcal{RF}$ alone. Instead, the waveforms in column C are almost inverted by the $\mathcal{IR}$ in the top two examples; In the bottom example, the fast temporal modulations in the input and in panel C are heavily filtered out by the $\mathcal{IR}$ to produce finally a smoother response (panel E).

The histogram in Figure 11 sums up the correlation coefficients found between predicted and measured responses for all units with tested with complex ripple combinations. As with simple ripple combinations, the majority of correlation coefficients are significantly large (89% with $\rho > 0.6$).

5 DISCUSSION

5.1 Summary of responses to moving ripple combinations

The results presented in this report 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, 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 (1995). Errors are also inevitable in

computing 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). Finally, measurements of the transfer functions and of the multiple ripple stimuli are done sequentially over a relatively long period of time (from 30 to 60 minutes) during which the state of the animal is likely to change somewhat.

5.2 The effects of response nonlinearities

AI units exhibit response nonlinearities due 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 by comparing the (linearly) predicted response waveforms (which can go negative) to 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 sharper (or having 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. 1995).

However, these largely instantaneous nonlinearities appear to act upon 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 hence "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.

5.3 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 to 16 Hz, with an approximate bandwidth of 3 octaves. These response parameters are summarized for all units isolated in Fig.6 in the companion paper (Kowalski et al. 1995). 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.

In the first view, the 8 Hz tuning may be seen as a general physiological limitation of cortical cell dynamics, or as an epi-phenomenon related to the rhythms induced by cortico-thalamic 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 abolish the sustained responses to stationary spectra and preserve 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 RF's. 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, 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 spectro-temporal decomposition is remarkably similar in spirit to the common practice in engineering systems (such as speech recognition systems) (Rabiner and Schafer 1979) 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.

6 Acknowledgments

This work is supported by grants from the Office of Naval Research, the Air Force Office of Scientific Research, a training grant from the National Institutes of Health (NIDCD T32 DC00046-01) and the National Science Foundation (#NSFD CD 8803012). We thank P. Gopalaswamy for help with the acquisition and analysis system, A. Owens, H. Versnel and T. Denison for assistance with animal preparation and data recordings. We also thank P. Cariani for insightful discussions.

References

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

[2] 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.

[3] 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.

[4] 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, NY: McGraw-Hill, 1972, p.283–340.

[5] Kowalski, N., Depireux, D.A., Shamma, S.A. Analysis of Dynamic Spectra in Ferret Primary Auditory Cortex: I. Characteristics of single unit responses to moving ripple spectra. *Institute for Systems Research Technical Report*, 1995.

[6] Rabiner, L.R. and Schafer, R.W. *Digital processing of speech signals.* Prentice-Hall, 1978.

[7] de Ribaupierre, F., Goldstein, M.H., and Yeni-Komshian, G. Cortical coding of repetitive acoustic pulses. *Brain. Res.* 48:205–225, 1972.

[8] Shamma, S.A., Fleshman, J.W., Wiser, P.R. and Versnel, H. Organization of response areas in ferret primary auditory cortex. *J. Neurophys.* 69:367–383, 1993.

[9] 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 Neuroscience*, 1:233–254, 1995.

[10] Shamma, S.A. and Versnel, H. Ripple analysis in ferret primary auditory cortex: II. prediction of unit responses to arbitrary spectral profiles. *Auditory Neuroscience*, 1:255–270, 1995.

[11] Wang, K. and Shamma, S.A. Spectral shape analysis in the central auditory system. *IEEE Trans. Speech Audio Proc.* 3:382–395, 1995.

1 List of Figures

Figure 1: Spectrograms of stimuli consisting of multiple moving ripple spectra. Each moving ripple is characterized by its ripple frequency (Ω), initial phase (Φ), and its velocity (ω).

(A): Spectrogram of a stimulus consisting of two moving ripples: $\Omega = 0.8$ cycles/octave, $\omega = 4$ Hz, $\Phi = \pi/6$ added to another ripple with $\Omega = 0.8$ cycles/octave, $\omega = 8$ Hz, and $\Phi = \pi$. The spectrogram of the resulting combination ripple is sinusoidal at every instant, with $\Omega = 0.8$ cycles/octave as illustrated by the upper cross-section. The time-course of each spectral component is demonstrated by the cross-section to the right, which is the sum of two sinusoidal waveforms at 4 and 8 Hz.

(B): Spectrogram of the addition of five ripples with random phases. The constituent ripples are: $\Omega = 0.4, 0.8, 1.2, 1.6, 2.0$ cyc/oct, $\omega = 4, 8, 12, 16, 20$ Hz, $\Phi = 0, 30, 41, -101, 25$ degrees.

Figure 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 $T_{\Omega_o}(\omega)$. The magnitude $|T_{\Omega_o}(\omega)|$ and unwrapped phase $\Phi_{\Omega_o}(\omega)$ of the transfer function are shown in the top plots. 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 in the bottom plot.

Figure 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 $T_{\omega_m}(\Omega)$. The magnitude $|T_{\omega_m}(\Omega)|$ and unwrapped phase $\Phi_{\omega_m}(\Omega)$ of the transfer function are shown in the top plots. 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 in the bottom plot.

Figure 4: Predicting the final response to multiple moving ripple stimuli.

(A): Spectrogram of the stimulus ($S(x,t)$), along with its ripple content (three moving ripples in this case, all at $\Phi = 0$). The gray-scale indicates relative amplitude of the spectrogram.

(B): $\mathcal{RF}$ of the cell (measured as described in Fig.3) with $BF = 3$ 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 which is the response of the unit due to the $\mathcal{RF}$ alone.

(D): The $\mathcal{IR}(\sqcup)$ 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): The 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. The dashed line is the zero spike count. All abscissas measure time in seconds; y-axes labeled on the left are normalized spike counts. Y-axes on the right side indicate actual spike count. Arrows indicate the location of $t = 0$ of the periodic functions in (C-E), relative to the corresponding period of the stimulus.

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

Figure 6: Examples of responses of a unit to three combination ripple stimuli: (top row) two ripples , (middle row) three ripples and (bottom row) four ripples. All details of the plots and computations are as in Fig.4.

Figure 7:

(A): Example of responses of a unit to two- and four-ripple stimuli. The unit has a strongly inverting $\mathcal{IR}$ function.

(B) Uncovering the half-wave rectified portion of the responses. The two stimuli illustrated are inverted relative to each other, and hence their responses are also inverted. The measured and predicted responses are indicated as described in Fig.4 before. The measured responses are also used to construct the "non-rectified" version of the response, depicted by the dashed curve in panel (E) of the bottom row. All other details of the plots and computations are as in Fig.4.

Figure 8: Histogram of the correlation coefficients between measured and predicted responses to stimuli with two, three or four moving ripples.

Figure 9: Examples of responses to complex combinations of ripples. Three complex combinations are presented to the same unit: (Top row) simulated FM stimulus composed of five ripples to produce effectively a single peak with velocity 20 oct/sec. (Middle row) temporal noise stimulus generated with fifteen random phase moving ripples all at $\Omega = 1.2$ cycles/octave. (Bottom row) ripple noise stimulus generated by adding five ripples with different Ω's and random phases, all moving at the same velocity. All other details of the plots and computations are as in Fig.4.

Figure 10: Responses of three different cells to three 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 five ripples. (Middle row) a simulated FM stimulus effectively traveling at 20 oct/sec. (Bottom row) a simulated FM stimulus effectively traveling at 50 oct/sec. All other details of the plots and computations are as in Fig.4.

Figure 11: Histogram of the correlation coefficients between measured and predicted responses to complex ripple stimuli with five or more moving ripples.

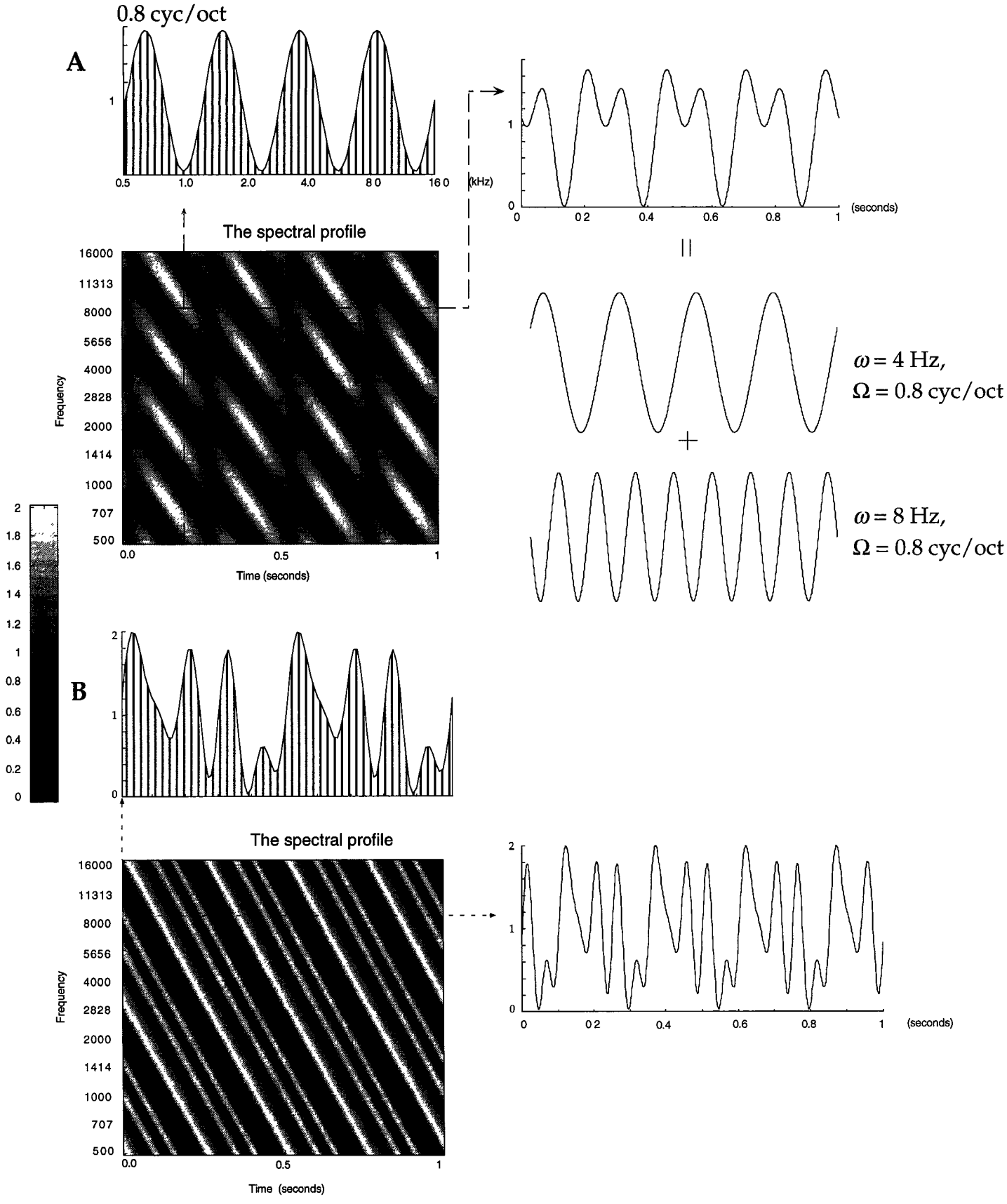

A
0.8 cyc/oct
1
0.5
1.0
2.0
4.0
8 0
16 0
(kHz)
The spectral profile
16000
11313
8000
5656
4000
2828
2000
1414
1000
707
500
Frequency
0.0
0.5
1
Time (seconds)
2
1.8
1.6
1 4
1.2
1
0.8
0.6
0.4
0.2
0
0 2
0.4
0.6
0.8
(seconds)
ω = 4 Hz,
Ω = 0.8 cyc/oct
ω = 8 Hz,
Ω = 0.8 cyc/oct
B
The spectral profile
Time (seconds)
(seconds)

Fig 1 - Vowels

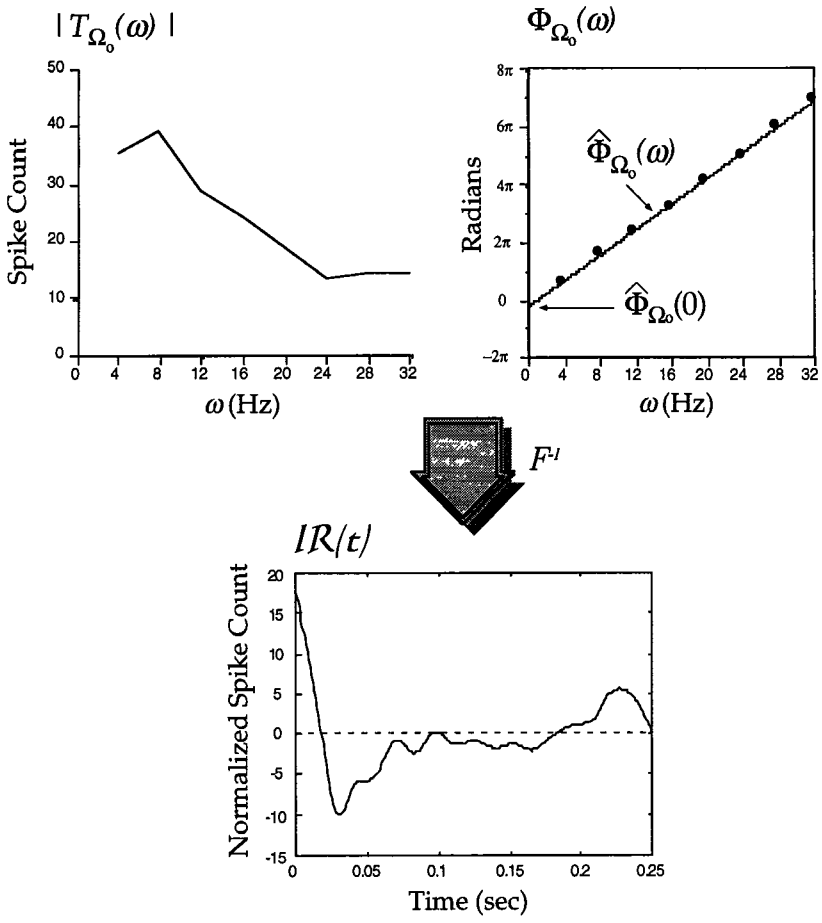

$|T_{\Omega_o}(\omega)|$
Spike Count
ω (Hz)
$\Phi_{\Omega_o}(\omega)$
Radians
$\hat{\Phi}_{\Omega_o}(\omega)$
$\hat{\Phi}_{\Omega_o}(0)$
ω (Hz)
F^{-1}
$IR(t)I$
Normalized Spike Count
Time (sec)

Fig 2 – Vow AL Sh

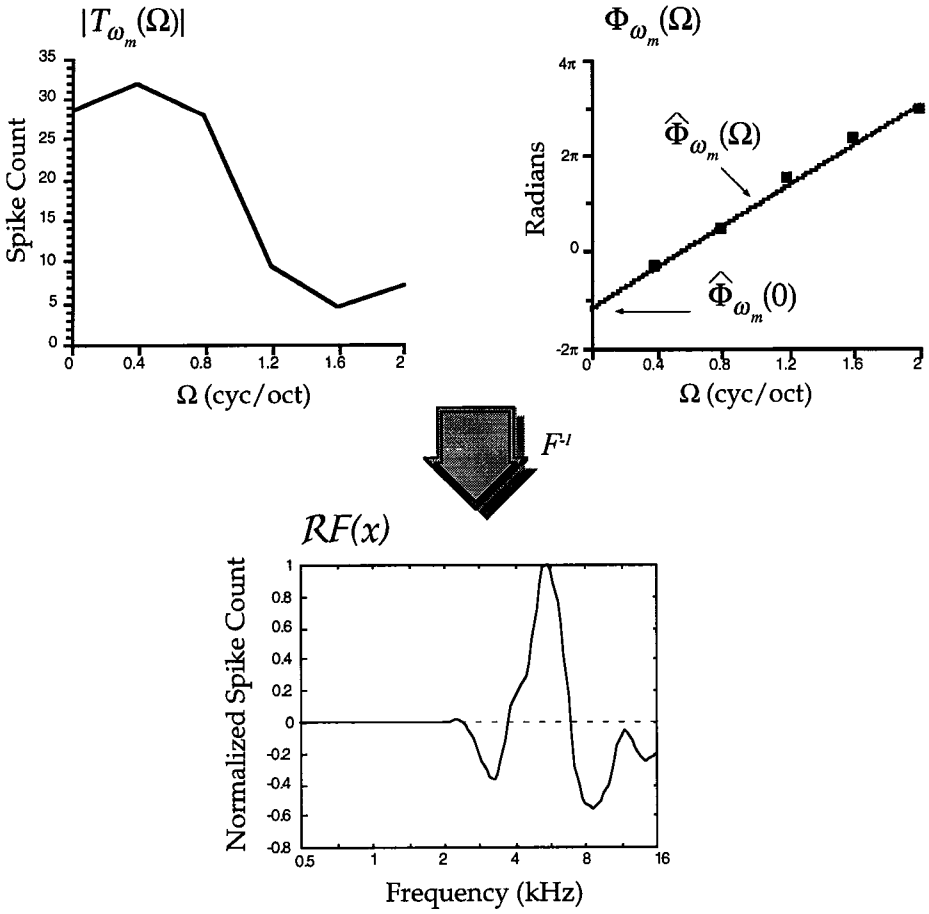

$|T_{\omega_m}(\Omega)|$
Spike Count
35
30
25
20
15
10
5
0
0
0.4
0.8
1.2
1.6
2
Ω (cyc/oct)
$\Phi_{\omega_m}(\Omega)$
Radians
4π
2π
0
-2π
$\hat{\Phi}_{\omega_m}(\Omega)$
$\hat{\Phi}_{\omega_m}(0)$
0
0.4
0.8
1.2
1.6
2
Ω (cyc/oct)
F^{-1}
$RF(x)$
Normalized Spike Count
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
0.5
1
2
4
8
16
Frequency (kHz)

Fig. 3 - Kowalski

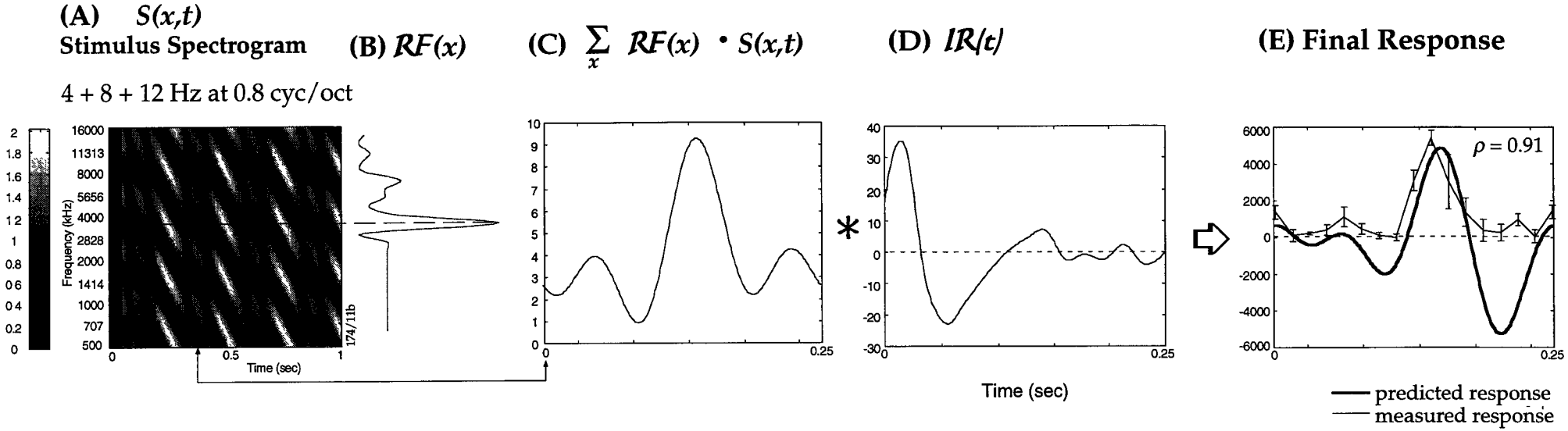
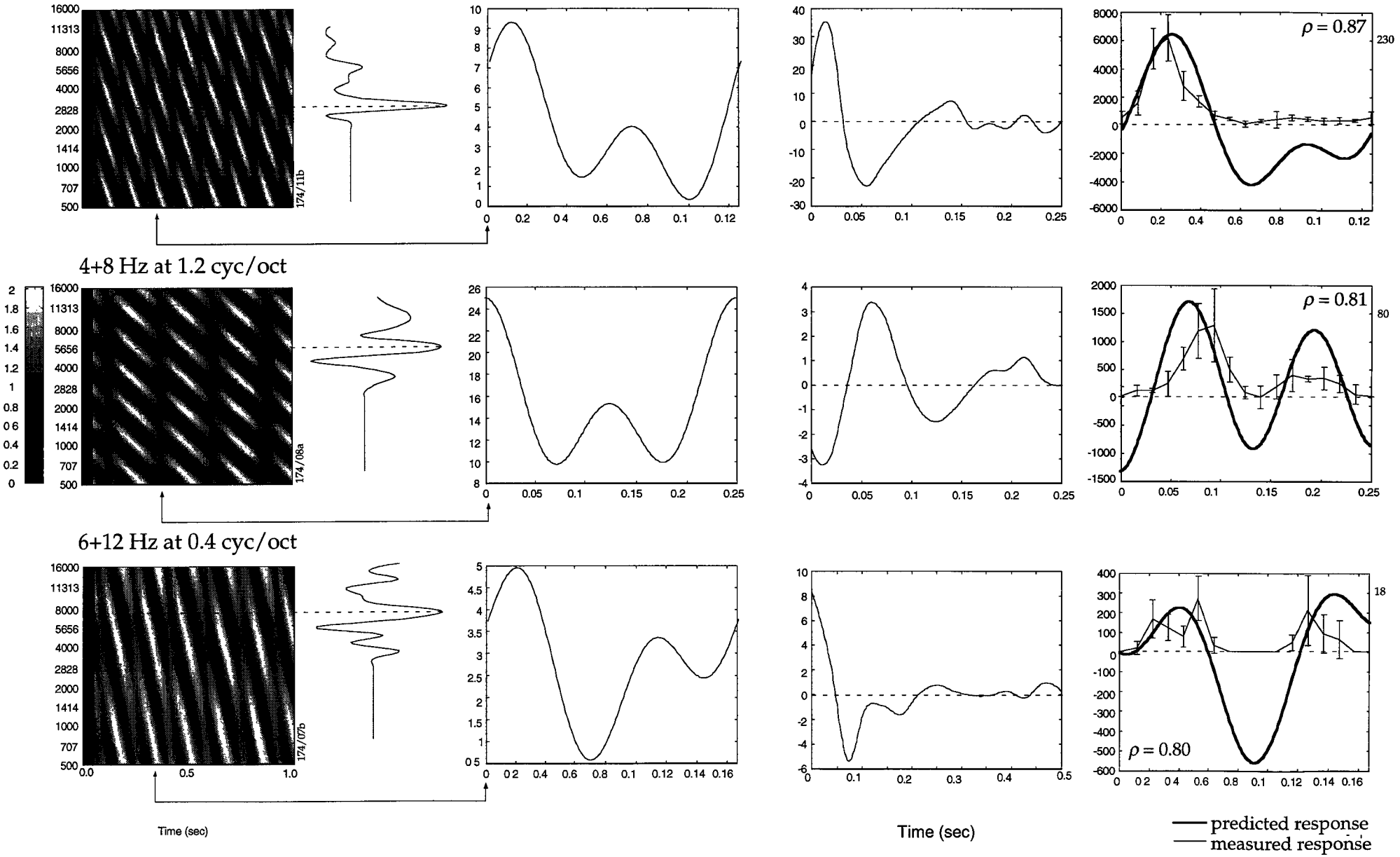

(A) S(x,t)
Stimulus Spectrogram
4 + 8 + 12 Hz at 0.8 cyc/oct
(B) RF(x)
(C) Σ RF(x) • S(x,t)
x
(D) IR(t)
(E) Final Response
Frequency (kHz)
Time (sec)
174/11b
ρ = 0.91
predicted response
measured response

Fig 4- KOWALSKI


(A) S(x,t)
Stimulus Spectrogram
(B) RF(x)
(C) Σ RF(x) • S(x,t)
x
(D) IR/t/
(E) Final Response
8 + 16 Hz at 0.8 cyc/oct
4+8 Hz at 1.2 cyc/oct
6+12 Hz at 0.4 cyc/oct
ρ = 0.87
ρ = 0.81
ρ = 0.80
Time (sec)
Time (sec)
predicted response
measured response

KOWALSKI

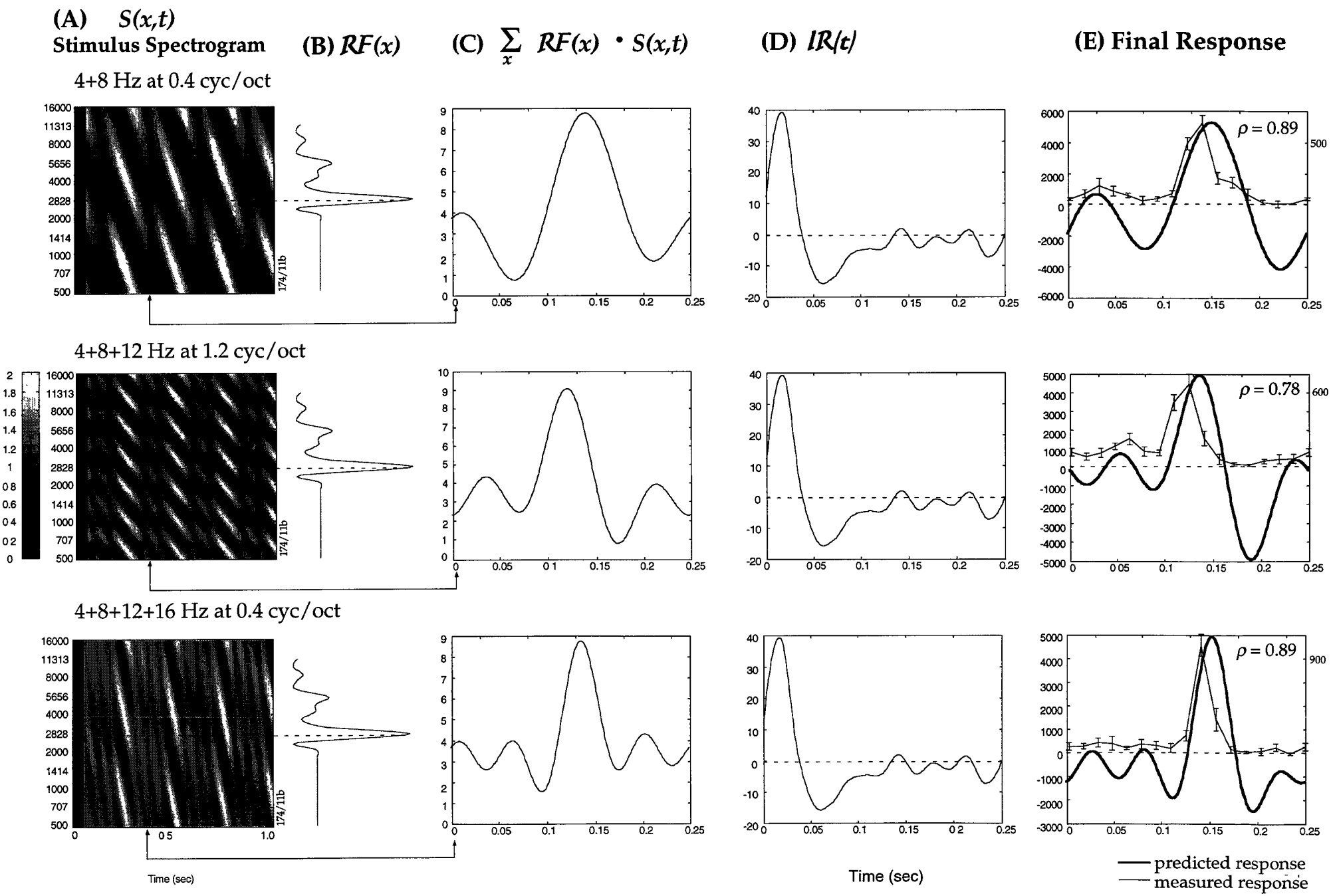

(A) S(x,t)
Stimulus Spectrogram
(B) RF(x)
(C) Σ RF(x) • S(x,t)
x
(D) IR/t/
(E) Final Response
4+8 Hz at 0.4 cyc/oct
4+8+12 Hz at 1.2 cyc/oct
4+8+12+16 Hz at 0.4 cyc/oct
16000
11313
8000
5656
4000
2828
2000
1414
1000
707
500
174/11b
ρ = 0.89
ρ = 0.78
ρ = 0.89
Time (sec)
predicted response
measured response

Fig 6. Kowalski

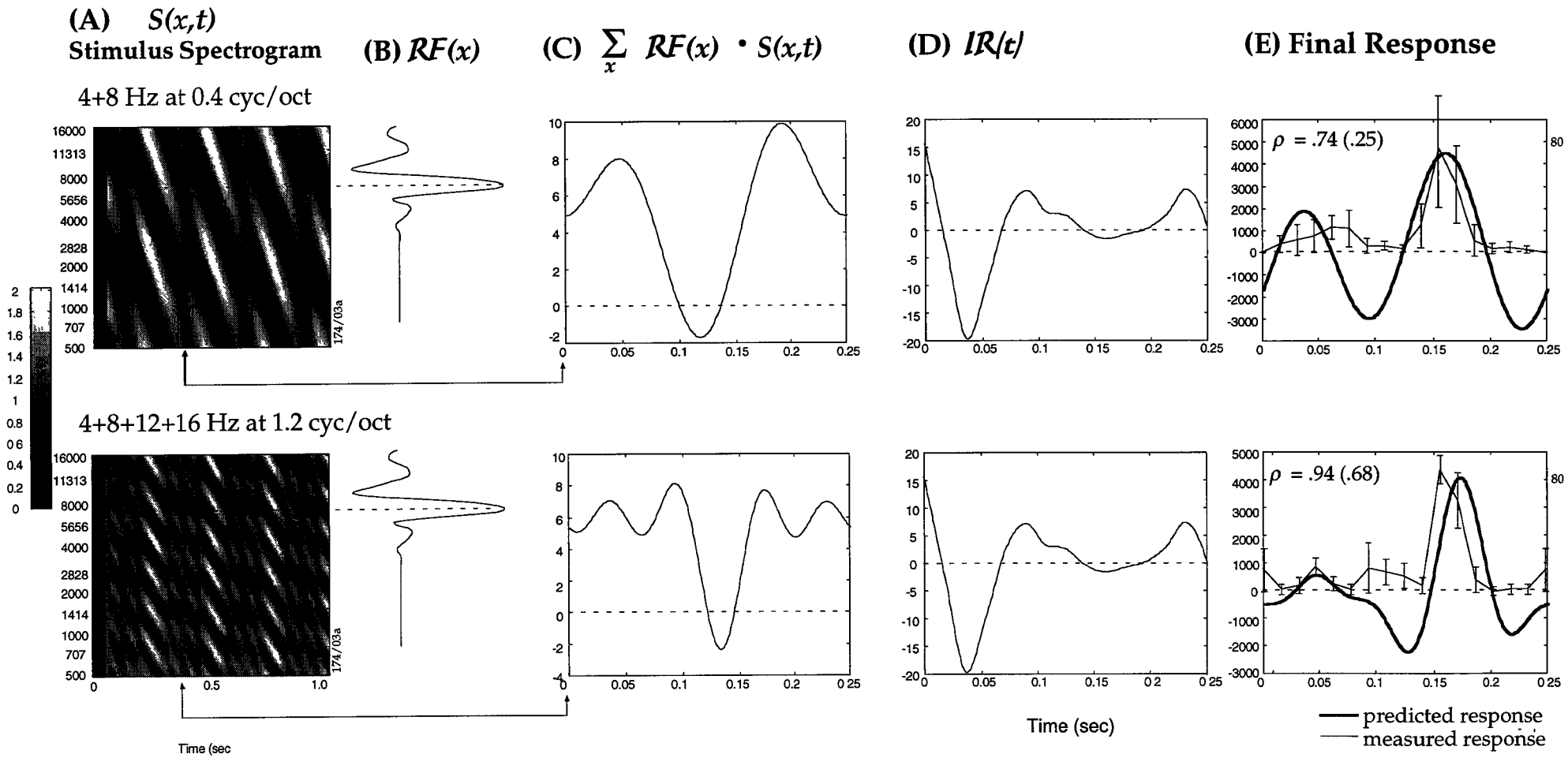

(A) S(x,t)
Stimulus Spectrogram
(B) RF(x)
(C) Σ RF(x) • S(x,t)
(D) IR/t/
(E) Final Response
4+8 Hz at 0.4 cyc/oct
4+8+12+16 Hz at 1.2 cyc/oct
16000
11313
8000
5656
4000
2828
2000
1414
1000
707
500
174/03a
Time (sec
ρ = .74 (.25)
ρ = .94 (.68)
Time (sec)
predicted response
measured response

Fig 7A. Kowalski

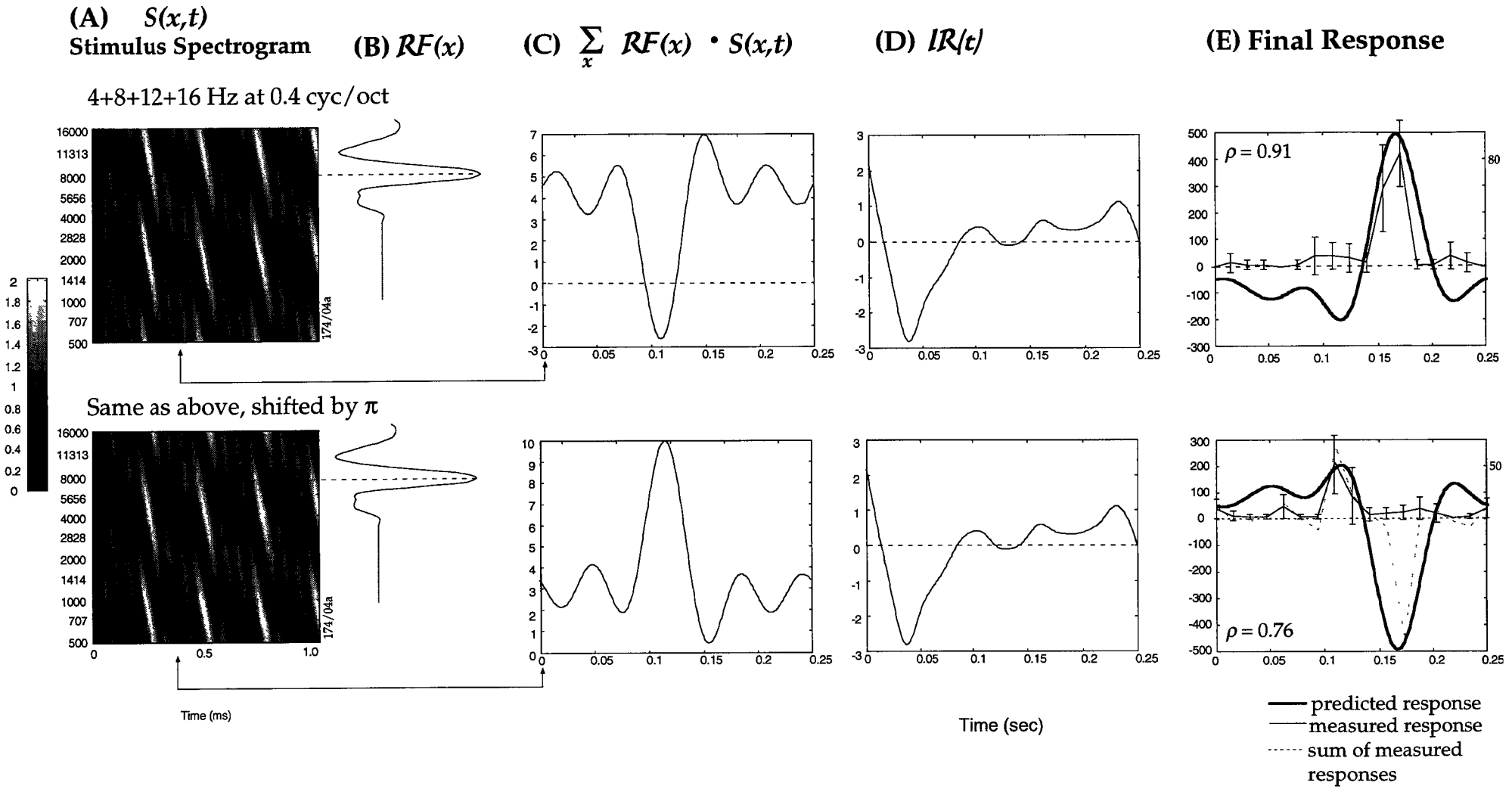

(A) S(x,t)
Stimulus Spectrogram
(B) RF(x)
(C) Σ RF(x) • S(x,t)
x
(D) IR/t/
(E) Final Response
4+8+12+16 Hz at 0.4 cyc/oct
Same as above, shifted by π
174/04a
Time (ms)
Time (sec)
ρ = 0.91
ρ = 0.76
predicted response
measured response
sum of measured responses

Fig 7 B Kawai et al.

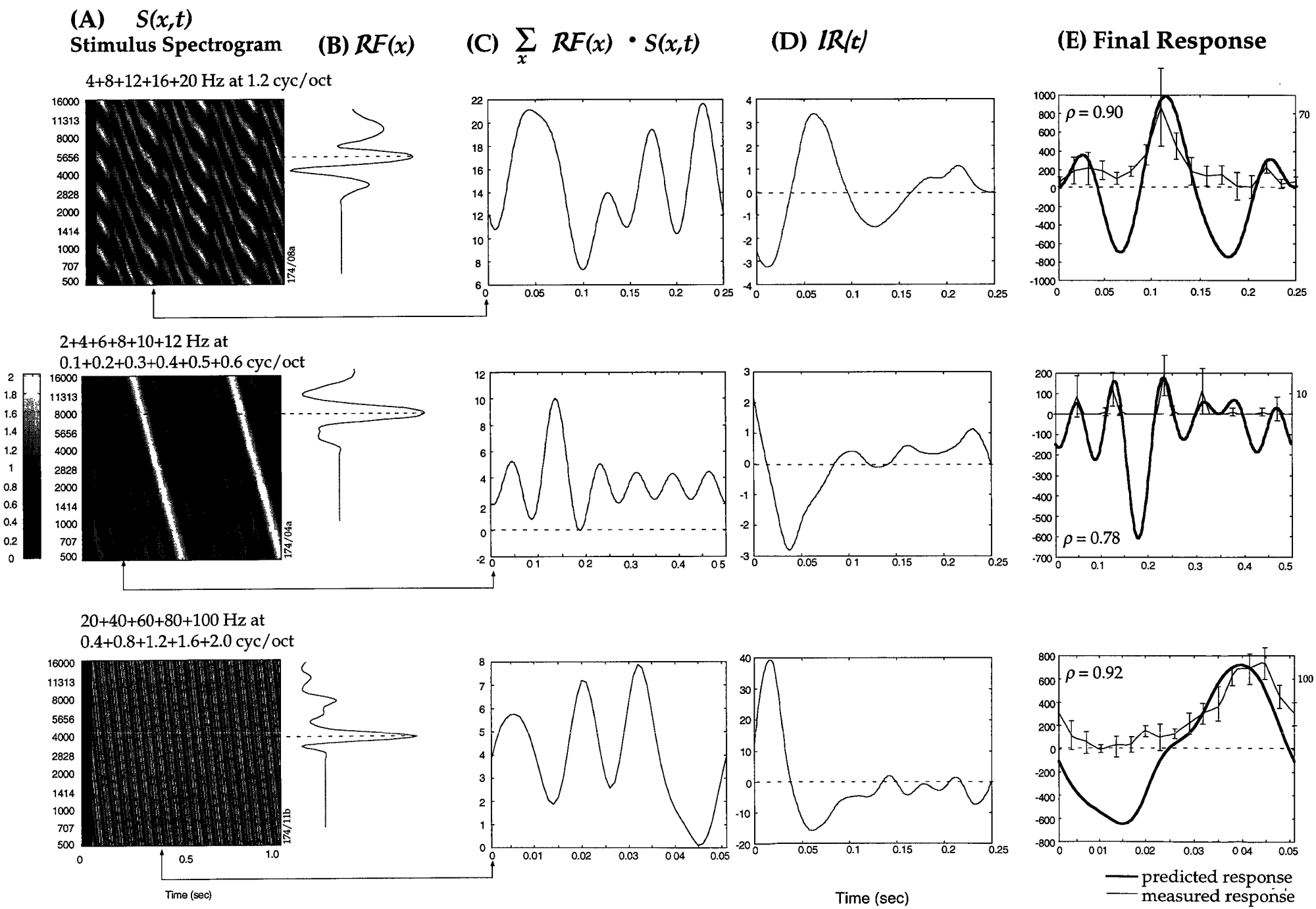

(A) S(x,t)
Stimulus Spectrogram
(B) RF(x)
(C) Σ RF(x) • S(x,t)
x
(D) IR(t)
(E) Final Response
4+8+12+16+20 Hz at 1.2 cyc/oct
2+4+6+8+10+12 Hz at
0.1+0.2+0.3+0.4+0.5+0.6 cyc/oct
20+40+60+80+100 Hz at
0.4+0.8+1.2+1.6+2.0 cyc/oct
174/08a
174/04a
174/11b
ρ = 0.90
ρ = 0.78
ρ = 0.92
Time (sec)
predicted response
measured response

Fig 9. Kowalski

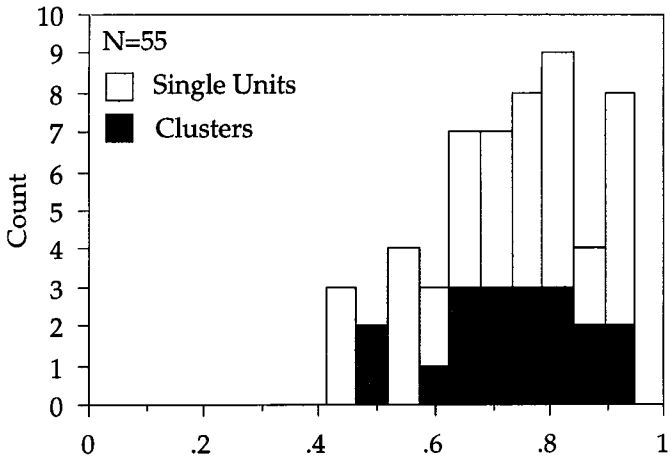

N=55
Single Units
Clusters
Count
10
9
8
7
6
5
4
3
2
1
0
0
.2
.4
.6
.8
1

Fig 8 - Kowalski

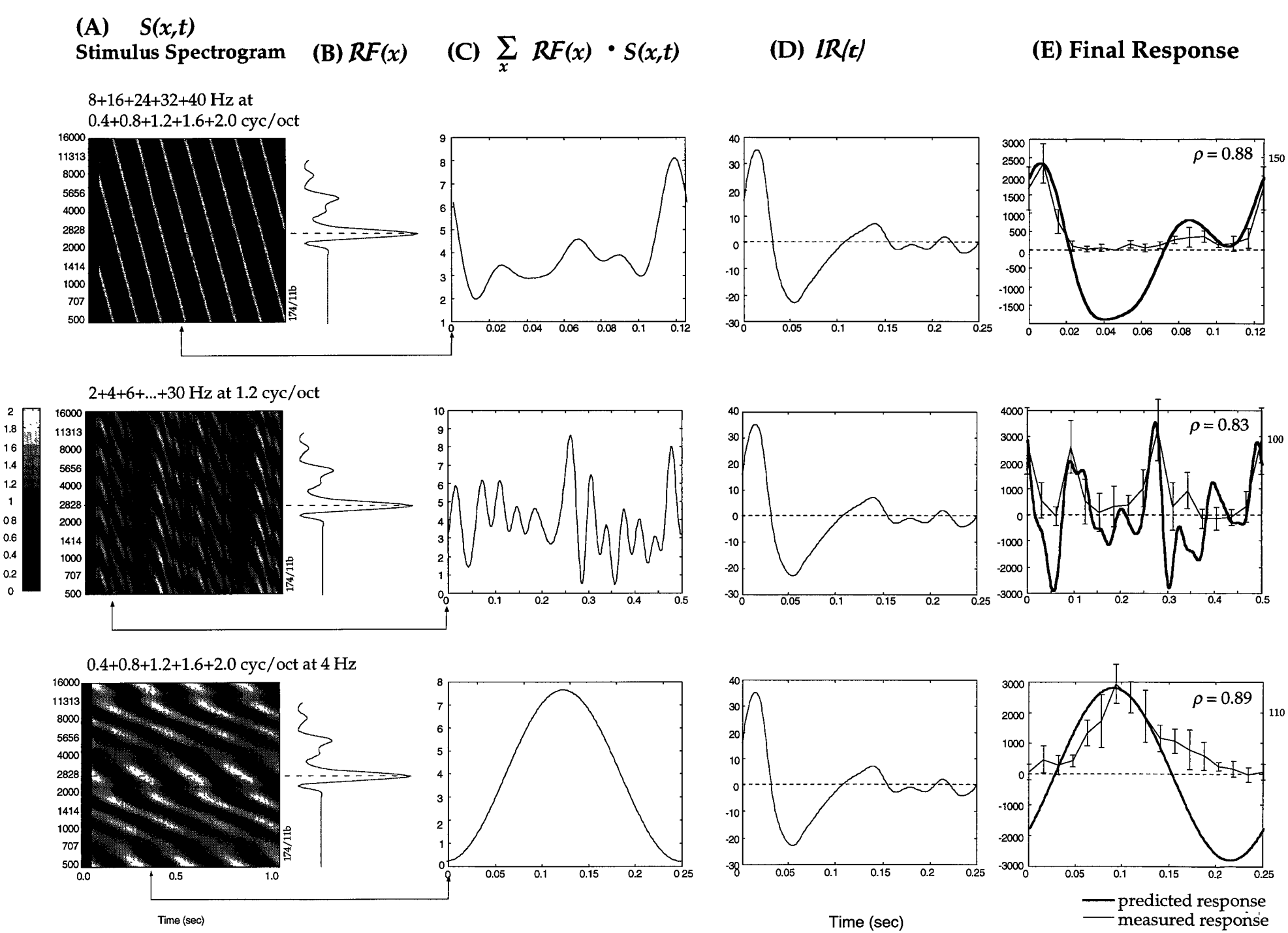

(A) S(x,t)
Stimulus Spectrogram
(B) RF(x)
(C) Σ RF(x) • S(x,t)
x
(D) IR(t)
(E) Final Response
8+16+24+32+40 Hz at
0.4+0.8+1.2+1.6+2.0 cyc/oct
2+4+6+...+30 Hz at 1.2 cyc/oct
0.4+0.8+1.2+1.6+2.0 cyc/oct at 4 Hz
ρ = 0.88
ρ = 0.83
ρ = 0.89
174/11b
Time (sec)
predicted response
measured response

Fig 10- KOWALSKI

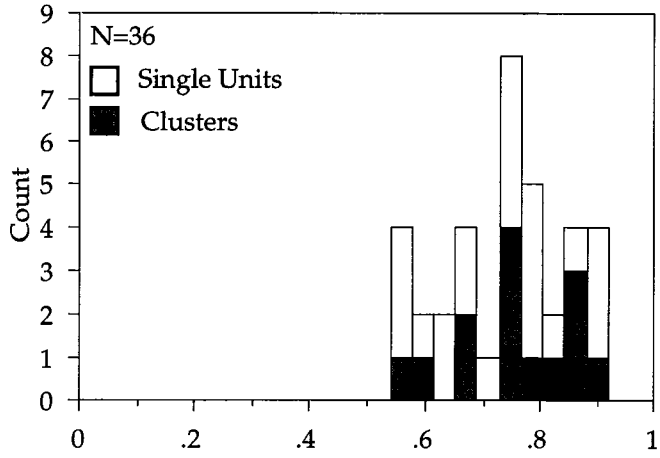

N=36
Single Units
Clusters
Count
9
8
7
6
5
4
3
2
1
0
0
.2
.4
.6
.8
1

Fig 11- KOWALSKI