Seventh Quarterly Progress Report

February 1 through April 30, 1994

NIH Contract N01-DC-2-2401

Speech Processors for Auditory Prostheses

Prepared by

Blake S. Wilson, Charles C. Finley, Mariangeli Zerbi and Dewey T. Lawson

Center for Auditory Prosthesis Research
Research Triangle Institute
Research Triangle Park, NC 27709

CONTENTS

I. Introduction

The purpose of this project is to design and evaluate speech processors for implantable auditory prostheses. Ideally, the processors will extract (or preserve) from speech those parameters that are essential for intelligibility and then appropriately represent those parameters for electrical stimulation of the auditory nerve or central auditory structures. Work in the present quarter included the following:

1. Studies with Ineraid subjects SR2 and SR3. The studies for both subjects included measures of intracochlear evoked potentials for a variety of stimuli. The studies with SR3 also included (a) measures of speech reception with *continuous interleaved sampling* (CIS) processors using different pulse durations and rates, (b) use of a forward masking technique to assess the spatial patterns of neural excitation produced by stimulation of single and multiple electrodes in the Ineraid implant, and (c) measures of speech reception with single-channel processors, for comparison with evoked potential and psychophysical results obtained at the University of Iowa (by C. J. Brown and P. J. Abbas) and in our laboratory for single electrodes with SR3's implant. The studies with SR2 also included further evaluation of an implant system proposed for use in China (see Appendix 1). This system includes a four-channel CIS processor and a four-channel transcutaneous transmission link, using four separate pairs of transmitting and receiving coils. In the studies with SR2, the outputs of the four passive receiving coils were routed to the implanted electrodes via the percutaneous connector of the Ineraid device.
2. Participation in a design review for a portable speech processor, under joint development by investigators at the Hôpital Cantonal Universitaire in Geneva, Switzerland, the Massachusetts Eye & Ear Infirmary in Boston, and RTI (design review held in Boston, February 16-18). The portable processor will be capable of implementing CIS and more complex processing strategies for implant patients outside of the laboratory. We expect that it will be ready for use in research studies by late summer, 1994.
3. Continued work with the Innsbruck team in helping to refine and evaluate their new portable processor. This included a design review at RTI in April, revision of certain aspects of the design following the review, and tests with Ineraid subjects SR3 and SR2 at Duke Medical Center in early May. In general, speech reception scores produced by the Innsbruck processor were quite similar to those produced by our laboratory system, using the same processor parameters.
4. Presentation of project results in invited lectures at the *Fifth Symposium on Cochlear Implants in Children* (New York City, February 4), the Indiana University School of Medicine (Indianapolis, March 9), and the *Annual Carolinas Audiology Conference* (Asheville, NC, April 1).
5. Continued preparation of manuscripts for publication.

In this report we describe initial results from recordings of intracochlear evoked potentials (EPs) for trains of pulses and for sinusoidally amplitude modulated (SAM) pulse trains. We also present a model of the population neural response, to these and other stimuli, developed by us under support of NIH project P01-DC00036, "Mechanisms of Intracochlear Electrical Stimulation." EP recordings and related psychophysical studies with human subjects were conducted under support of the present project. Results from other studies indicated above will be presented in future reports.

II. Temporal Representations with Cochlear Implants

Much of our prior work has been directed at improving spatial representations with cochlear implants, e.g., through elimination of simultaneous channel interactions with CIS processors [Wilson et al., 1991] and through controlled shifts in sites of neural excitation with "virtual channel" CIS (VCIS) processors [Wilson et al., 1994]. We now are directing a large fraction of our effort to investigation of temporal representations. This has included (a) development of a model to predict the population response of the auditory nerve to electrical stimulation, (b) recordings of intracochlear evoked potentials for repetitive stimuli, and (c) measurement of psychophysical responses to the same or similar stimuli. In general, results to date point to crude and highly limited representations. Several repair strategies have emerged from the modeling studies, however, that may produce improvements in speech reception for implant patients. The purpose of this report is to describe the model and to present initial findings from the related EP and psychophysical studies with human subjects.

Modeling Studies

A central consideration in the design of cochlear prostheses is how to represent speech sounds in the patterns of electrically evoked discharges in the auditory nerve. Although results of single-unit studies help to characterize the responses of individual neurons to electrical stimuli [Clopton and Glass, 1984; Dynes and Delgutte, 1992; Glass, 1985; Hartmann et al., 1984; Hartmann and Klinke, 1990; Javel, 1990; Javel et al., 1987; Kiang and Moxon, 1972; Knauth et al., 1994; Parkins, 1989; Parkins and Colombo, 1987; van den Honert and Stypulkowski, 1984 and 1987a,b], the input received by the central auditory nervous system (CANS) for suprathreshold stimuli is an ensemble of responses from a population of excited neurons. Therefore, perception with cochlear implants is linked to patterns of ensemble responses. Improved control of percepts elicited by cochlear prostheses -- and improved representations of speech and other signals -- almost certainly will require an understanding of (a) the biophysical mechanisms that underlie ensemble responses and (b) the interpretation of the peripheral input by the CANS.

Design and Construction of Ensemble Models

Models of the ensemble response of the auditory nerve to intracochlear electrical stimulation can be constructed by coupling mathematical descriptions of electric field patterns in the cochlea with mathematical descriptions of neural responses elicited by electrical stimuli. Examples of these components of ensemble models are presented in Table 1. The simplest ensemble model couples an exponential falloff description of the electric field with a strength-duration description for prediction of neural responses to the imposed electric field. This combined model can provide a useful tool for demonstrating basic patterns of neural responses evoked by transient stimuli such as single pulses [Wilson et al., 1985; Wilson, 1986; Wilson and Finley, 1986].

An important objective of recent work has been to develop an ensemble model capable of predicting responses to repetitive or continuous stimuli. The form of the current model is illustrated in Figs. 1 and 2. As shown in Fig. 1, up to 40 neurons can be specified at each of 504 cochlear positions. The input for the neurons at any one position is the electric field strength at that position. In general, field

Table 1. Components of ensemble models.

Complexity	Electric field description	Neural response description
1	exponential falloffs	strength-duration curves
2	finite-element model of fields produced by various configurations of electrodes, in an isotropic, homogeneous medium	Hill model [1936] of neural membrane, coupled with a description of refractory effects
3	torus version of above, to describe effects of cross-turn and multiple-channel stimulation	simple incorporation of stochastic properties, e.g., using Verveen's modification [1962; Verveen and Derksen, 1968] of the Hill model
4	complete finite-element models of fields produced in the anisotropic, heterogeneous structures of the cat and human cochleas	modified McNeal model [1976], in which (a) membrane parameters appropriate for auditory nerve fibers are substituted for the parameters of much-larger, uniform-diameter fibers, (b) finite impedances of myelin segments are represented, and (c) node dynamics are altered to reflect those of mammalian fibers

strengths will vary across positions according to the type and placement of intracochlear electrodes.

In the simplest model for each of the neurons, a Hill-type membrane is used [Hill, 1936], with refraction and an optional source of membrane or synaptic noise (Fig. 2). Rheobase, chronaxie, and refractory properties are specified for each neuron. The software includes facilities for highly flexible specification of electric field shape and stimulus parameters. Standard conditions used for the majority of modeling studies to date have included (a) simulation of a "sharp" monopolar electric field, with a space constant of 3.6 mm; (b) a 3.6-to-1 linear distribution of rheobases, starting at 15 μA and ending at 54 μA; (c) a chronaxie of 400 μs for each neuron; (d) an absolute refractory period of 700 μs for each neuron; and (e) a time constant of 1250 μs for the relative refractory period of each neuron.

An example of model inputs and outputs is presented in Fig. 3. The top panel shows the time waveform of a stimulus presented to a monopolar electrode at the 12 mm position along the length of the scala tympani. The spatial distribution of the electric field during the excitatory phase of each stimulus pulse is shown in the left panel. The space constant for the exponential falloff with increasing distance from the stimulating electrode is 3.6 mm. Responses of single neurons are indicated in the central panel, which shows a dot diagram of the times and locations of predicted spike occurrences.

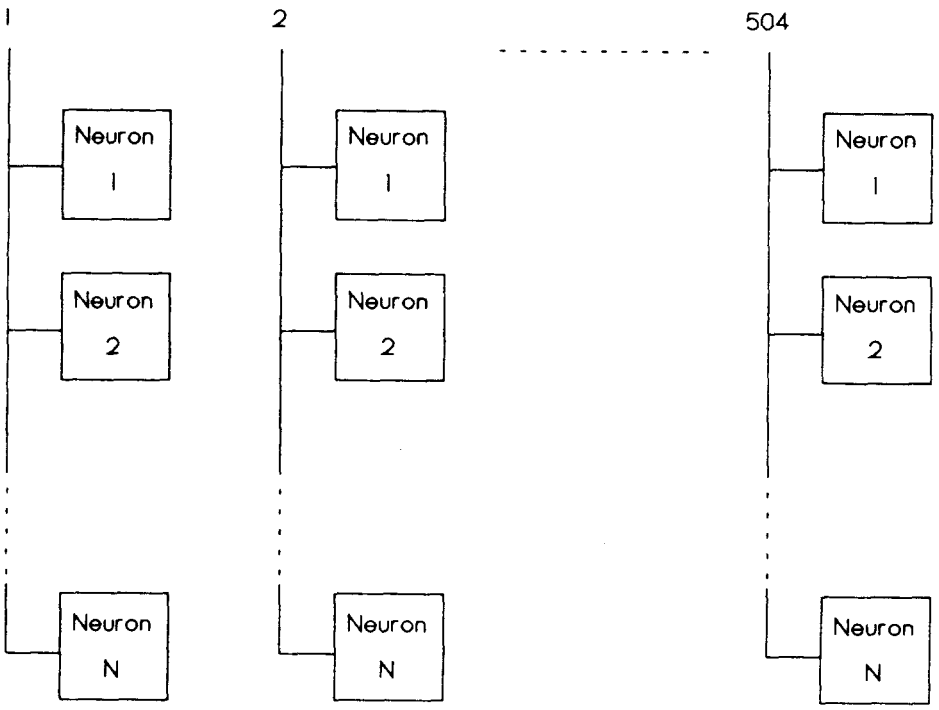

Fig. 1. Schematic diagram of an ensemble model.

The ordinate is location along the cochlea and the abscissa is time after stimulus onset (or latency). The dots can be coded with a color scale or gray scale (illustrated) to indicate the number of spike occurrences at a given position on the location-time map. The right panel shows a profile of the total number of spikes versus position integrated over time. It indicates the spatial pattern and extent of the response. The bottom panel shows total spike counts versus time, integrated over the spatial dimension. It provides an indication of how well the auditory nerve as a whole may follow or represent the waveform of the stimulus. In this case, such following is not particularly good, with a large neural response to pulse 1, no response to pulse 2, and a complicated pattern of responses thereafter.

Predicted Patterns of Response for Pulse Trains, Sinusoids, and SAM Pulse Trains

A sampling of results from model simulations using the standard conditions is presented in Figs. 4 through 8. Fig. 4 shows predicted patterns of response for pulse trains, sinusoids, and sinusoidally amplitude modulated (SAM) pulse trains, for various choices of rates, frequencies or modulation frequencies, respectively. Note that in the predictions the nerve follows well the rates or frequencies of stimulation up to 200 pps or 200 Hz, but fails to reflect fully the stimulus waveform at higher rates or frequencies. An alternating response is observed, for example, for stimulation with pulses presented at 300 pps, where relatively large responses are elicited by the first and subsequent odd-numbered pulses,

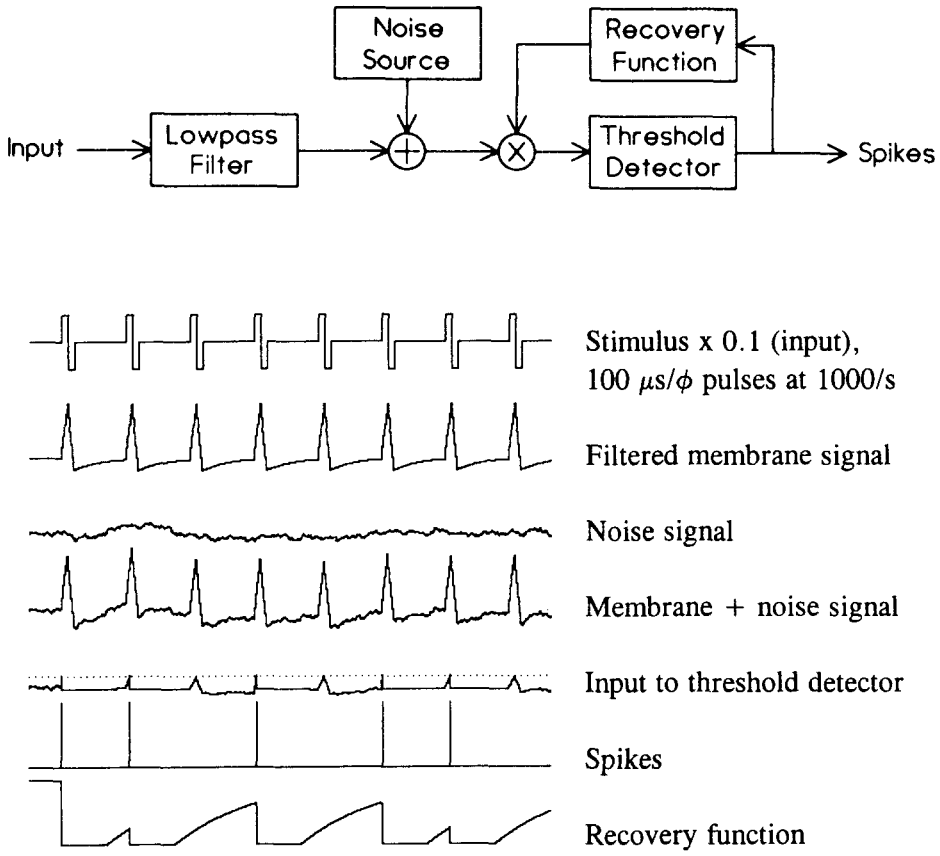

Fig. 2. Modified Hill model of a single neuron. The input (electric field strength outside the neuron) is filtered to simulate the lowpass characteristic of passive neural membranes. The corner frequency of this first-order filter is set to reflect the specified chronaxie of the neuron. An optional source of noise may be added to the membrane signal. (The spectrum of the noise can be altered with simple filters, or shaped to approximate other functions, such as the 1/f spectrum described by Verveen [1962] and Verveen and Derksen [1968]. In the majority of present simulations involving the addition of noise a 200 Hz lowpass filter has been used, as illustrated here.) The product of the membrane plus noise signal and a recovery function is compared with a threshold, corresponding to the specified rheobase for the neuron. Times of neural discharge (spikes) are indicated with each positive crossing of the threshold. Periods of absolute and relative refraction post discharge are simulated with a recovery function of the form:

$$\text{Recovery} = 0, \text{ for } t < a, \text{ and}$$
$$\text{Recovery} = 1 - e^{-(t-a)/\tau}, \text{ for } t \geq a,$$

where t is time post discharge, a is the period of absolute refraction, and τ is the time constant of relative refraction.

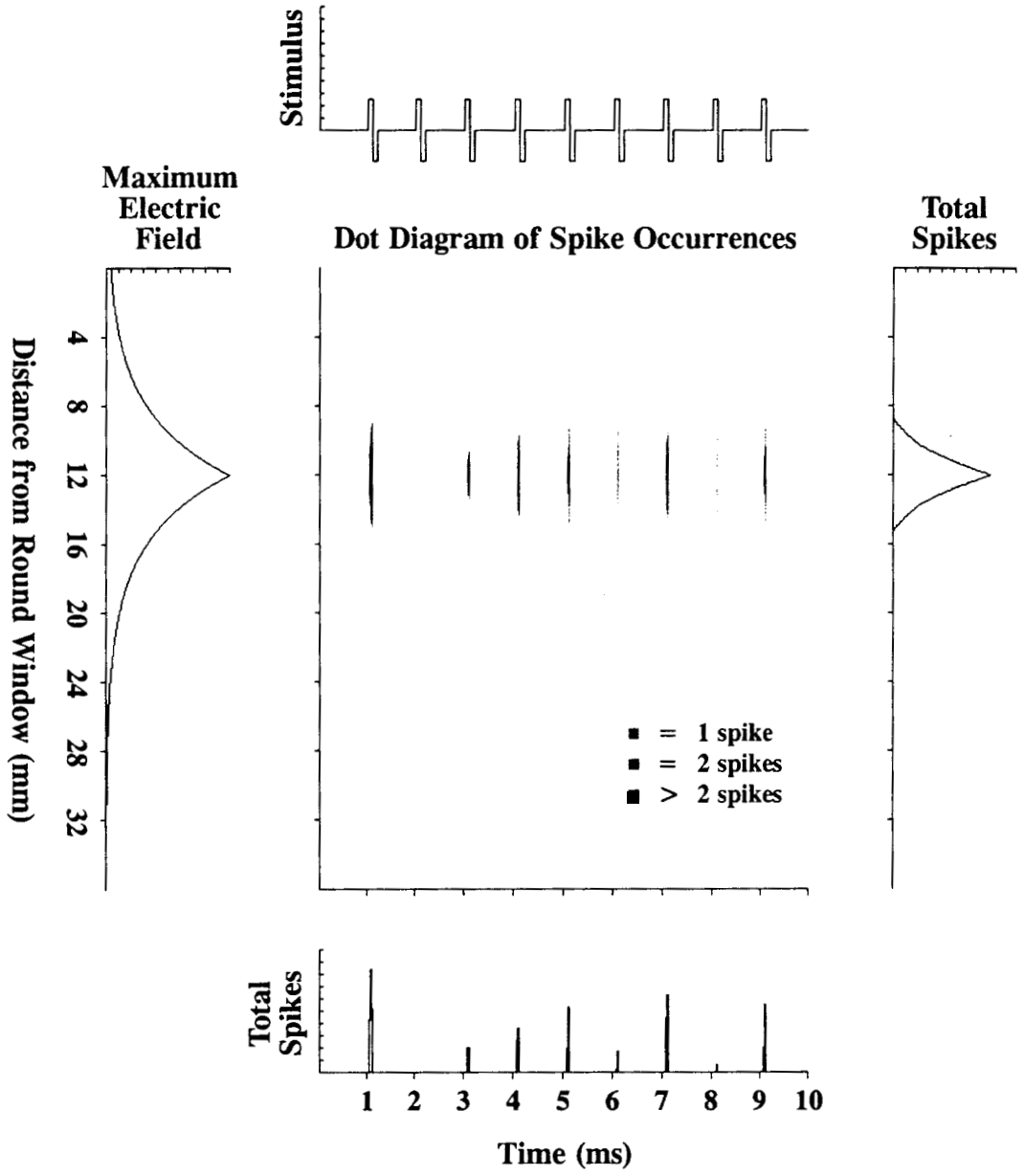

Fig. 3. Example of model inputs and outputs. Inputs include specification of the stimulus waveform, site(s) of the electrode(s) used to deliver the stimulus, falloff of the electric field with increasing distance from each site of stimulation, and parameters for each of the neurons. The stimulus here consisted of a train of 100 μs/phase pulses presented at the rate of 1000/s (top panel). The amplitude of each phase was 180 μA. The space constant of the electric field was 3.6 mm (as illustrated in the left panel). All neurons had a chronaxie of 400 μs, an absolute refractory period of 700 μs, and a time constant of 1250 μs for the subsequent period of relative refraction. At each spatial position, the 40 neurons had a 3.6-to-1 linear distribution of rheobase values, starting at 15 μA. Outputs of the model include a dot diagram of predicted spike occurrences (center panel), a spatial integration of spike occurrences (right panel), and a temporal integration of spike occurrences (bottom panel).

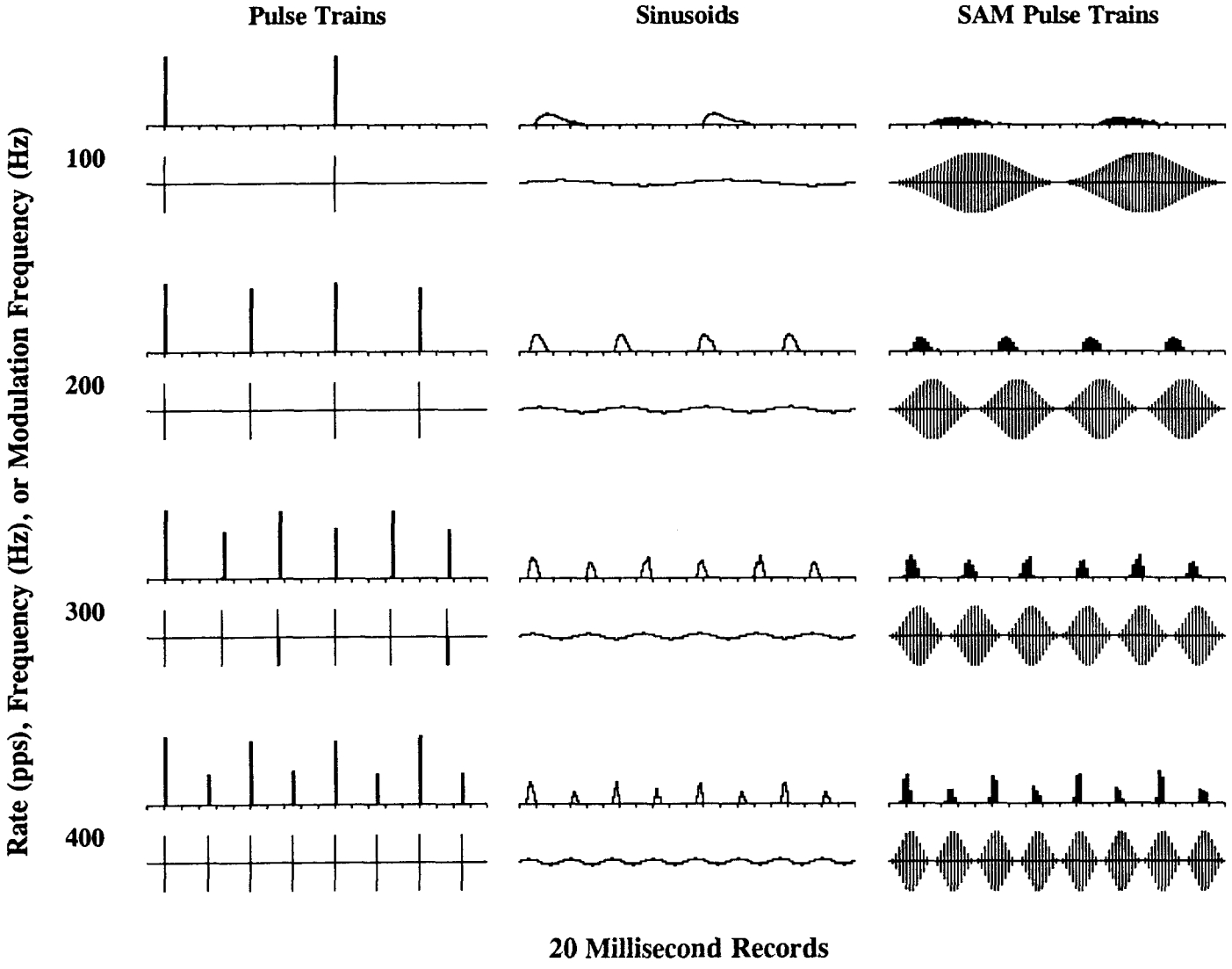

Fig. 4. Predicted patterns of spike counts versus time for the aggregate of modeled neurons, for various stimuli. In each panel, the stimulus waveform is presented below the predicted pattern of neural responses (corresponding to the bottom panel in Fig. 3). The time axes are graduated in 1 ms steps, for a total duration of 20 ms in each case. The stimuli include trains of 30 μs/phase pulses presented at the indicated rates (left column), sinusoids with the indicated frequencies (middle column), and sinusoidally amplitude modulated (SAM) pulse trains with the indicated modulation frequencies (right column). The carrier for the SAM pulse trains consists of 30 μs/phase pulses presented at 5000 pps. The depth of modulation is 100 percent. Amplitudes for the three types of stimuli were adjusted to produce approximately equal numbers of spikes across the types at a given rate or frequency (i.e., across rows in the figure). Note the broad similarity of responses to sinusoids and SAM pulse trains for each of the four frequencies (responses to SAM pulse trains appear to be "filled in" because the responses are a concatenation of responses to individual pulses in the carrier, whereas the responses to sinusoidal stimuli reflect a continuous distribution of activity during the excitatory phases of the stimulus waveforms).

and smaller responses by the second and subsequent even-numbered pulses. This appears to be an expression of refraction in the nerve, with all neurons available for response to the first pulse and with many neurons -- still in relative refraction at that time -- therefore unavailable for stimulation by the second pulse. This interplay between stimulus rate and refraction is further illustrated in Fig. 5, which shows predicted patterns of response for three pulse amplitudes and for a wider range of pulse rates.

Predicted patterns of response to a wide variety of SAM pulse trains are presented in Figs. 6 and 7. The patterns indicate that representation of the modulation waveform can be quite crude or distorted, particularly for low carrier rates. The representation of 300 Hz modulation for the 1000 pps carrier rate, for example (Fig. 7), shows a variety of intervals, not one of which corresponds to the reciprocal of the modulation frequency. Also, the conditions for the carrier rate of 250 pps show clearly the effects of aliasing in the representation, i.e., stimuli and responses are identical for modulation frequencies of 100 and 150 Hz, and for 50 and 200 Hz (Fig. 6).

Influence of Membrane or Synaptic Noise on the Predicted Response

The presence of membrane or synaptic noise can have a strong effect on predicted patterns of response by imparting a degree of stochastic independence among neurons. This effect is illustrated in Fig. 8, where various levels of noise have been added to the membrane signal in the neural models (see Fig. 2), holding all other conditions constant. As is obvious from the figure, the ability of the aggregate population to follow pulses presented at 1000 pps is improved with increasing levels of noise up to level 3 or 4. Levels 2 and higher produce spontaneous discharges in the absence of stimulation (as may be seen prior to the first pulses in Fig. 8). The average rate of spontaneous activity produced by level 3 corresponds closely to the average rate found in normal hearing, presumably produced by continuous release of chemical transmitter substance into the synaptic clefts between hair cells and neurons.

Psychophysical and Electrophysiological Studies

Such predictions of ensemble responses have been evaluated in psychophysical and electrophysiological studies with human subjects. The psychophysical studies have included measures of stimulus scaling and identification, and the electrophysiological studies have included measures of intracochlear evoked potentials, obtained in recordings from unstimulated electrodes in the Ineraid implant array. Stimuli have included 200 ms bursts of pulses presented to different electrodes at various rates (100 to 1000 pps, in steps of 100 pps) and levels (at most comfortable loudness and at several 10% steps down in amplitude from that corresponding to MCL). Stimuli also have included SAM pulse trains, with carrier rates of 1000, 500 and 250 pps, and with modulation frequencies of 50, 100, 150 and 200 Hz for the two lower carrier rates and 50, 100, 150, 200, 300 and 400 Hz for the 1000 pps carrier rate. Three subjects have participated in the studies to date.

Measures of intracochlear evoked potentials for subject SR2 are presented in Figs. 9, 12, 15 and 16. In all cases electrode 3 of the Ineraid implant was used for stimulation (with respect to a remote electrode in the temporalis muscle) and potentials were recorded differentially between intracochlear electrode 4 and an electrode at the ipsilateral mastoid. Body potential was measured with a reference electrode at the wrist.

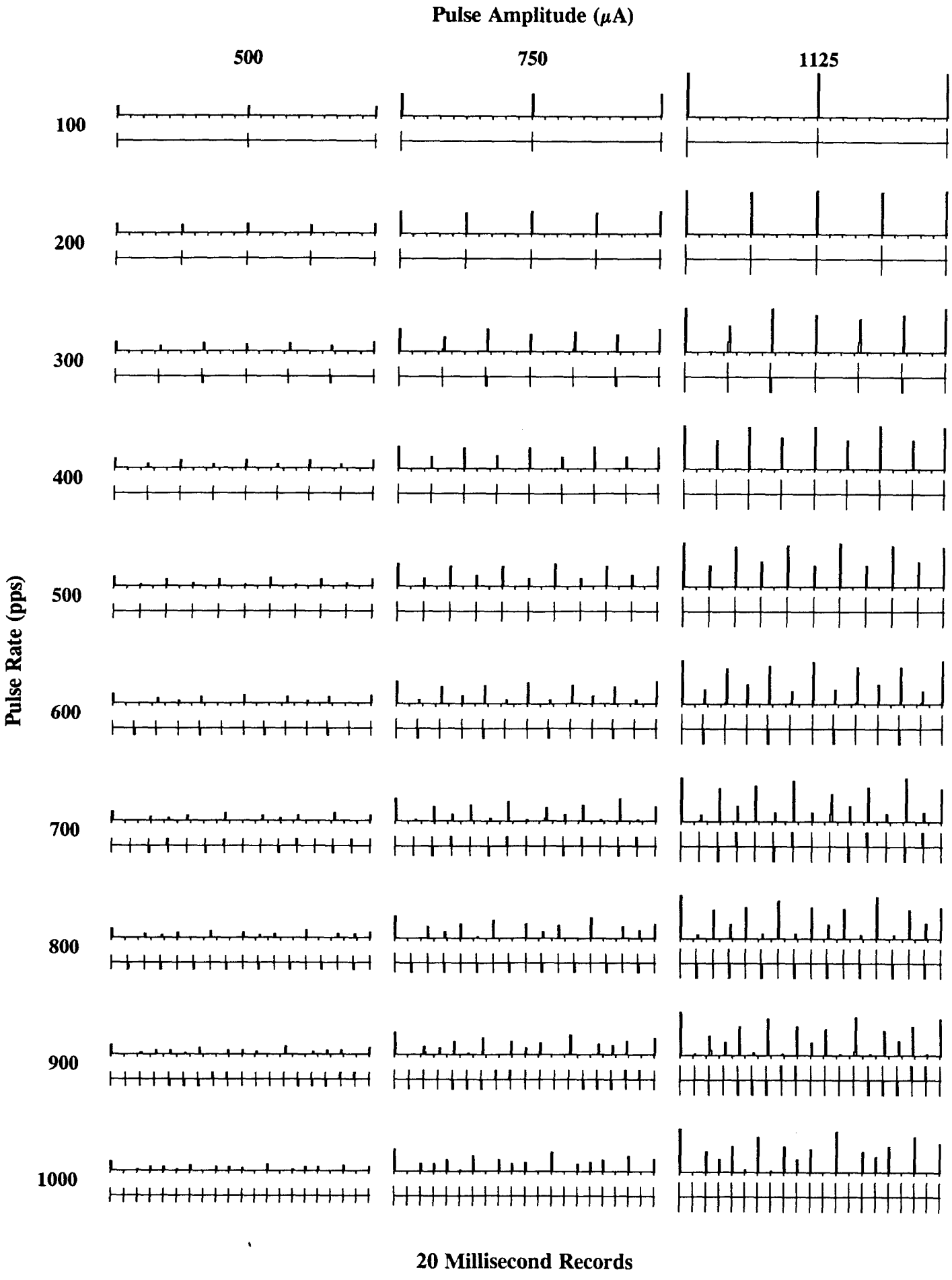

Fig. 5. (please see caption on next page)

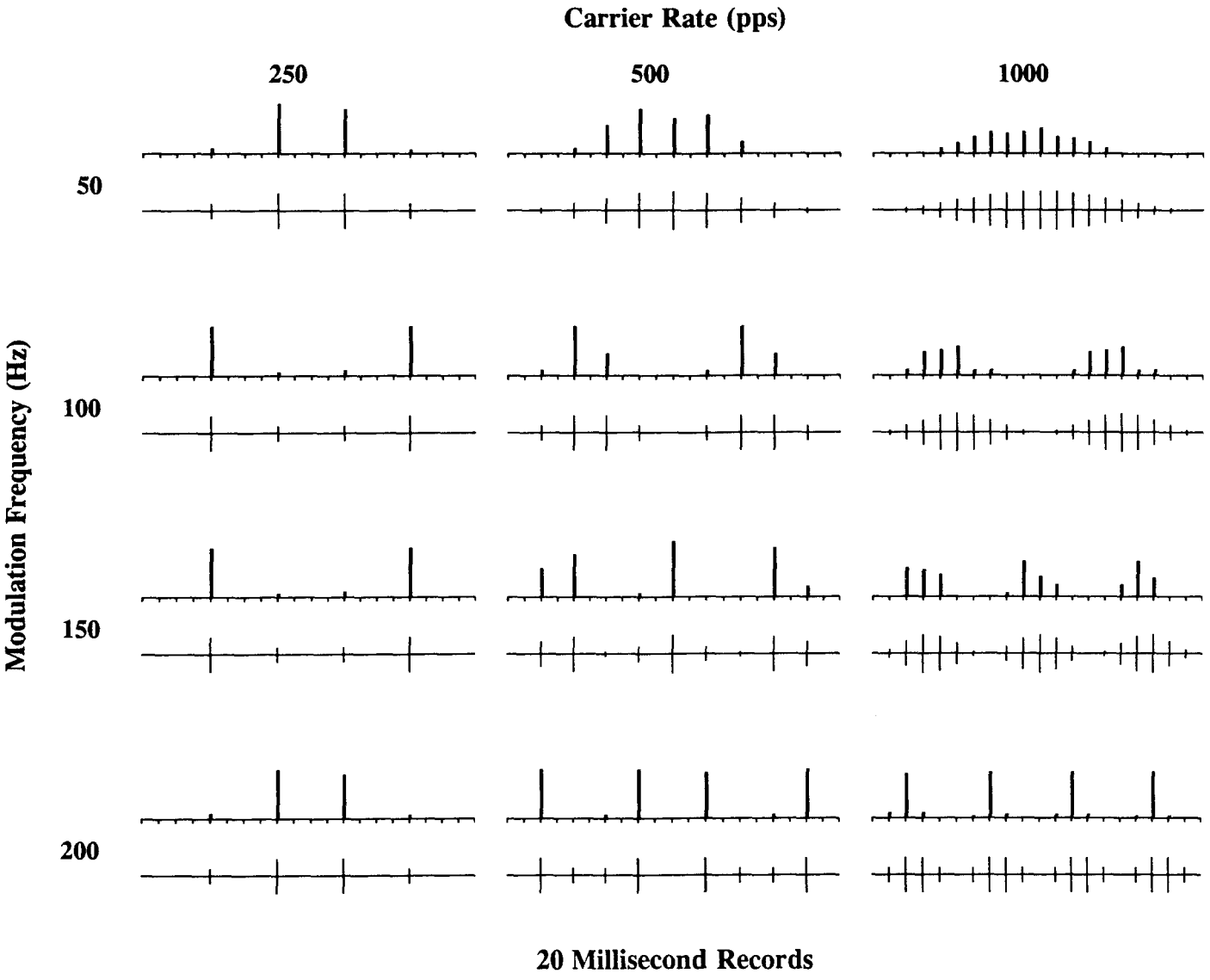

Fig. 6. Predicted patterns of spike counts versus time for the aggregate of modeled neurons, for sinusoidally amplitude modulated (SAM) pulse trains with the indicated modulation frequencies and carrier rates. The duration of pulses in the carriers was 30 μs/phase, and the depth of modulation 100 percent. Organization of each panel is the same as in Fig. 4.

Fig. 5. (from the prior page) Predicted patterns of spikes counts versus time for the aggregate of modeled neurons, for trains of 30 μs/phase pulses presented at the indicated rates and amplitudes. As in Fig. 4, each panel shows the stimulus waveform below the predicted pattern of neural responses. The time axes are the same as those in Fig. 4. Note the progressive degradation in the representation of the stimulus waveform as the pulse rate is increased beyond 200/s. Note also the subtle differences in the representations across pulse amplitudes, especially at the higher pulse rates.

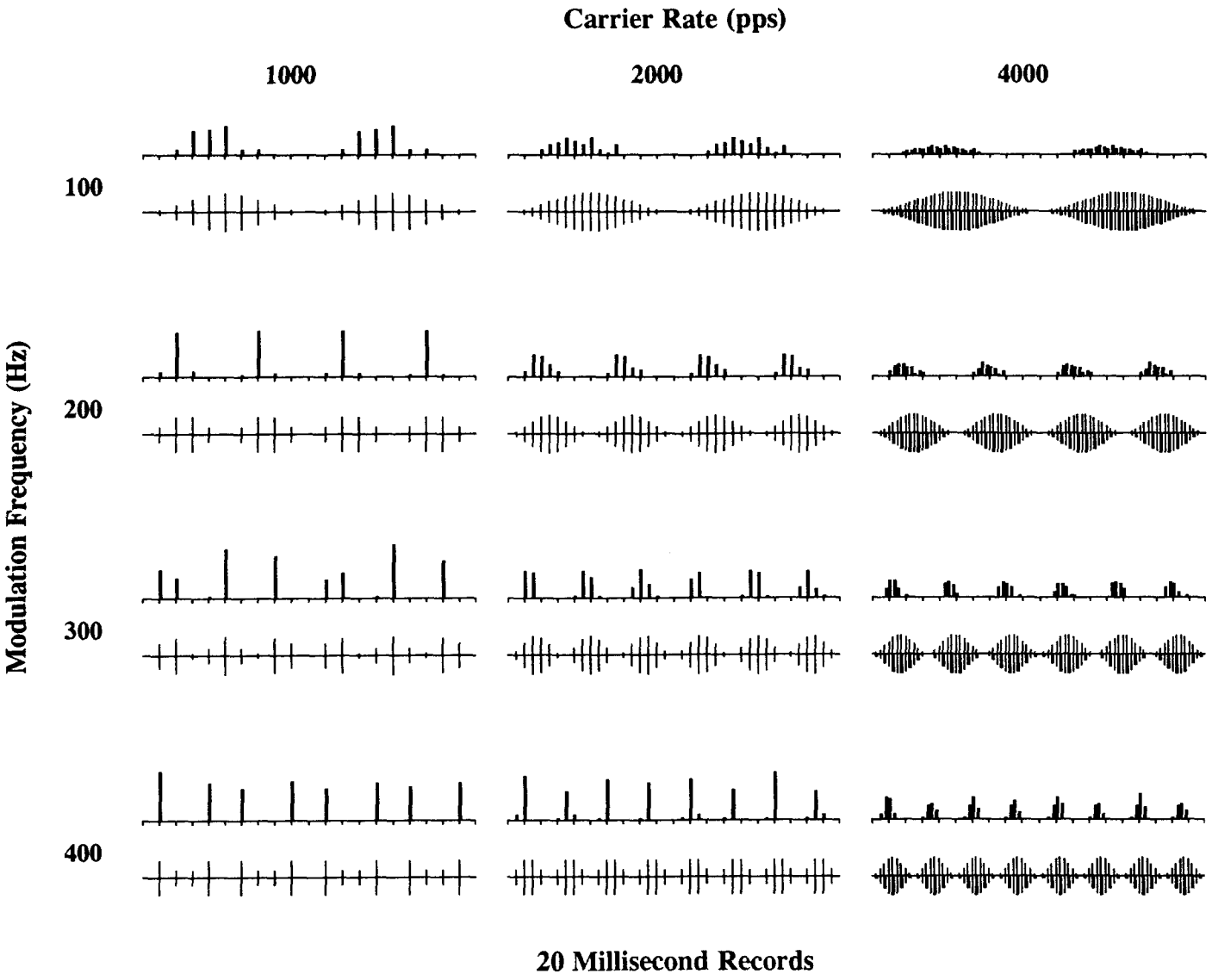

Fig. 7. Predicted patterns of spike counts versus time for the aggregate of modeled neurons, for sinusoidally amplitude modulated (SAM) pulse trains with the indicated modulation frequencies and carrier rates. The duration of pulses in the carriers was 30 μs/phase, and the depth of modulation 100 percent. Organization of each panel is the same as in Fig. 4. This series includes higher carrier rates and a wider range of modulation frequencies than shown in Fig. 6.

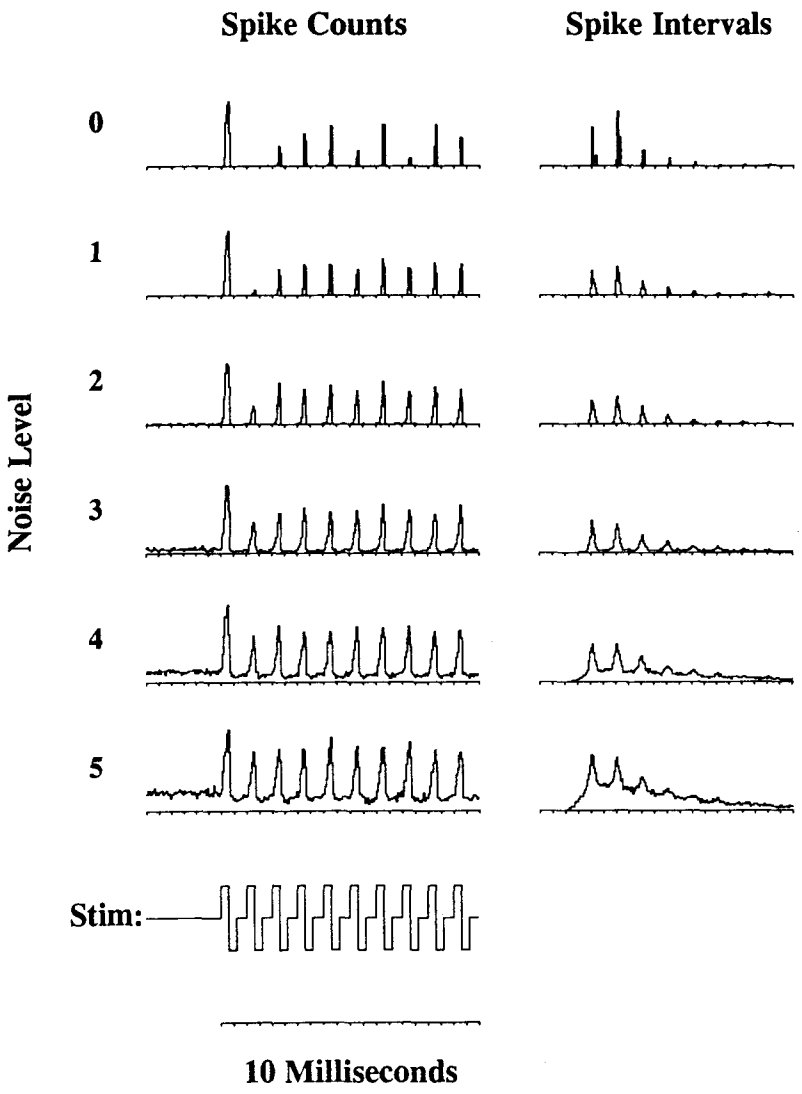

Fig. 8. Predicted patterns of spike counts versus time, and intervals between spikes for fibers with multiple responses, for the aggregate of modeled neurons. For each row in the figure, a different level of membrane noise (200 Hz lowpass noise, as in Fig. 2) was added for each of the neurons in the simulation, ranging from no noise (level 0, top row) to levels of noise capable of eliciting considerable amounts of spontaneous activity in the absence of stimulation. The stimulus waveform, shown beneath the Spike Counts column, consisted of 300 μs/phase pulses presented at 1000 pps. Note the improvement in the neural representation of the stimulus waveform with increasing levels of membrane noise up to level 3 or 4. Note also that even a low level of membrane noise, below the level required for eliciting spontaneous activity in the absence of stimulation, can have a profound effect on the neural representation of the stimulus waveform (see predicted pattern of responses for Noise Level 1). Time axes are calibrated in 500 μs intervals.

Responses to Trains of Pulses

Fig. 9 shows responses to pulses presented at various rates. The EPs following each pulse have equal magnitudes, presumably reflecting equal numbers of single-unit discharges, for the pulse rates of 100 and 200/s. At higher rates, however, the EPs show a failure of the whole-nerve response to reflect fully the waveforms of the repetitive stimuli. For rates of 400 pps and higher an alternating pattern of response is observed, consistent with model predictions (see Fig. 4, left column).

As has been described by Brown and Abbas [1990a; 1990b], a recovery function can be derived from the ratios of EP magnitudes for pairs of pulses, where the time between pulses is manipulated. A recovery function derived from such ratios for the first two pulses in the present results is presented in Fig. 10. The form of EP waveshapes, and the derived recovery function, are quite similar to those of Brown and Abbas, for their patients with excellent recovery from prior stimulation and relatively large EPs (present EPs are approximately 80 $\mu V_{p\text{-}p}$ for the first pulse).

Note that although the patterns of EPs for pulse trains are generally consistent with model predictions, the predictions presented in Fig. 5 indicate a marked departure from a simple alternating pattern of responses for pulse rates above 600-700/s. For the higher rates of stimulation the present model predicts a more complex pattern of response than is observed in SR2's EPs. Note also that the values of absolute refraction and the time constant of relative refraction derived from the data of Fig. 10 differ somewhat from the values used in our standard model.

These results suggest that substitution of the values indicated in Fig. 10 might produce closer agreement between model and data in comparing, e.g., Figs. 5 and 9. Also, the addition of a low level of membrane noise might improve agreement between model and data, in that the data indicate a somewhat better ability of the nerve to follow repetitive stimuli at high rates than predicted by the model, using the standard conditions.

Additional aspects of whole-nerve responses to trains of pulses are illustrated in Fig. 11, which shows derived EP magnitudes for the full 200 ms of the bursts used in the experiment of Fig. 9. First, note that EP magnitudes are relatively uniform across pulses for the 100 pps condition. The $\pm 10\%$ variability in magnitudes may reflect variability in the measurement with a limited number of sweeps. For the 200 pps condition, a small decrement (approximately 20%) is observed in the magnitudes of EPs over the 200 ms burst. Presumably this is a reflection of accommodation or adaptation in the response. A similar decrement in the response is observed for higher rates of stimulation after the first two or three pulses. The apparent accommodation at 1000 pps is not obviously different from that at 200 pps. For the rates of 400 pps and higher the response to the second pulse in the burst is substantially lower than the response to the first pulse. Responses to subsequent pulses show the alternating pattern described above for some period into the burst. A small alternation in the response may be present for the first several pulses following the second pulse for the 400 pps condition (i.e., roughly 10 ms into the burst). A clear alternation persists over the first 30 ms or so for the 600 and 800 pps conditions, and over the first 80 to 100 ms for the 1000 pps condition. The depth of the alternating response appears to increase with increases in rate of stimulation.

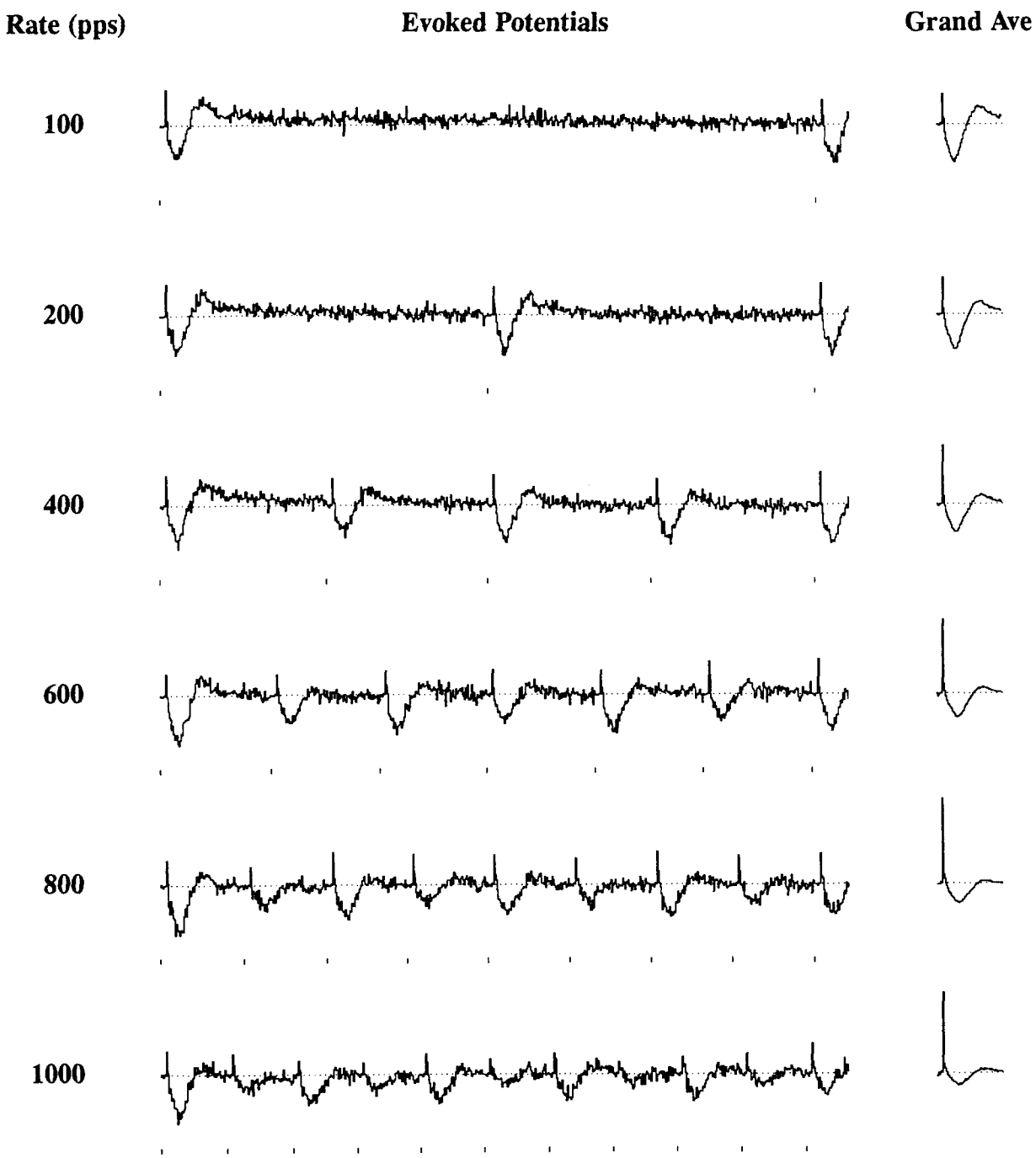

Fig. 9. Recordings of intracochlear evoked potentials for Ineraid subject SR2 (please see full caption on next page).

Fig. 9. (from the prior page) Recordings of intracochlear evoked potentials for Ineraid subject SR2. Stimuli included 16.4 μs/phase pulses presented at the indicated rates to electrode 3 in the Ineraid array (the array has 6 intracochlear electrodes, with electrode 1 the most apical). Pulse amplitude was 750 μA for all illustrated conditions. This amplitude produced a most comfortable loudness (MCL) percept for the pulse rate of 1000/s. Lower pulse rates at the same amplitude produced lower loudnesses. The times of pulse presentations are indicated in the figure with the short vertical lines beneath each EP trace. Potentials were recorded differentially between intracochlear electrode 4 and the ipsilateral mastoid. Body potential was measured with a reference electrode at the wrist. A blanker circuit was used during pulse presentations to reduce the magnitude of pulse artifacts in the recordings. In addition, responses from 200 sweeps of pulses with the positive phase leading were added to responses from 200 sweeps of pulses with the negative phase leading. As illustrated, these two procedures reduced pulse artifacts to a low level (residual artifacts are seen in the "spikes" preceding each EP). The average of EPs following each pulse for a given condition is shown in the right column, under the heading of "Grand Ave." The horizontal dotted lines in the EP columns indicate zero potential. Note that EPs fail to follow pulses with equal amplitudes for rates of stimulation above 200 pps.

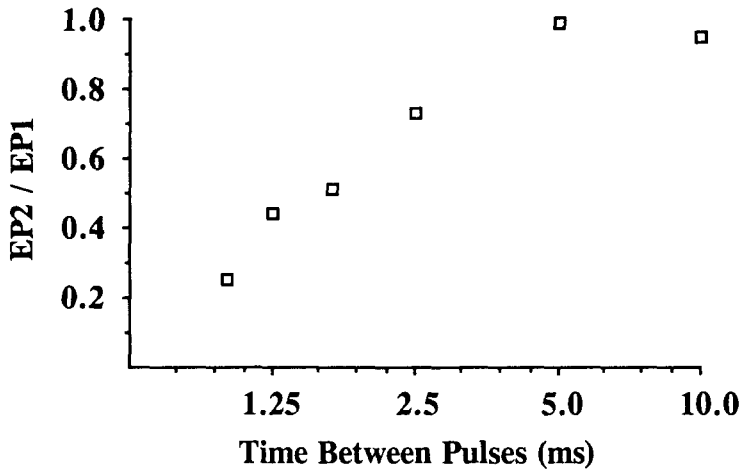

Fig. 10. Ratio of the amplitude of the evoked potential produced by the second pulse for each condition in Fig. 8 to that produced by the first pulse. Amplitudes were measured from N_1 to P_1. Note full recovery of EP amplitude for times between sequential pulses of 5 ms or greater (corresponding to pulse rates of 200 pps or less). For shorter periods between pulses, recovery is not complete. The amount of recovery (i.e., the EP2 / EP1 ratio) for those shorter periods can be fit by a general recovery equation of the form:

$$EP2/EP1 = 1 - e^{-(t - a)/\tau},$$

where t is time between pulses, a is a period of absolute refraction, and τ is the time constant of relative refraction. A least squares fit of the equation to the first four points in the figure ($r = 0.99$) indicates an absolute refractory period of 0.49 ms and a time constant of 1.55 ms for the period of relative refraction. These values are somewhat different than the values of our standard conditions for Figs. 3 through 8 ($a = 0.70$ ms; $\tau = 1.25$ ms). Substitution of the values indicated by the recovery function in the present figure may produce closer agreement between model and data in comparing, e.g., Figs. 5 and 9.

Noise in the recordings may mask these features to some extent. A recording of the response to pulses at 1000 pps obtained with 1000 sweeps for the two pulse polarities is presented in Fig. 12, and EP magnitudes derived from that recording are shown in the bottom panel of Fig. 13. Note that the level of noise in the recordings is reduced substantially by increasing the number of sweeps from 200 to 1000 (compare top and bottom panels in Fig. 12). Also note that, for the 1000 sweeps recording, the alternating pattern of response now is observed over the entire 200 ms of the burst (Fig. 13). It appears that an alternating response may not be not be visible in the EP data when the difference between sequential high and low responses is comparable to the level of noise in the recordings. Thus, when the

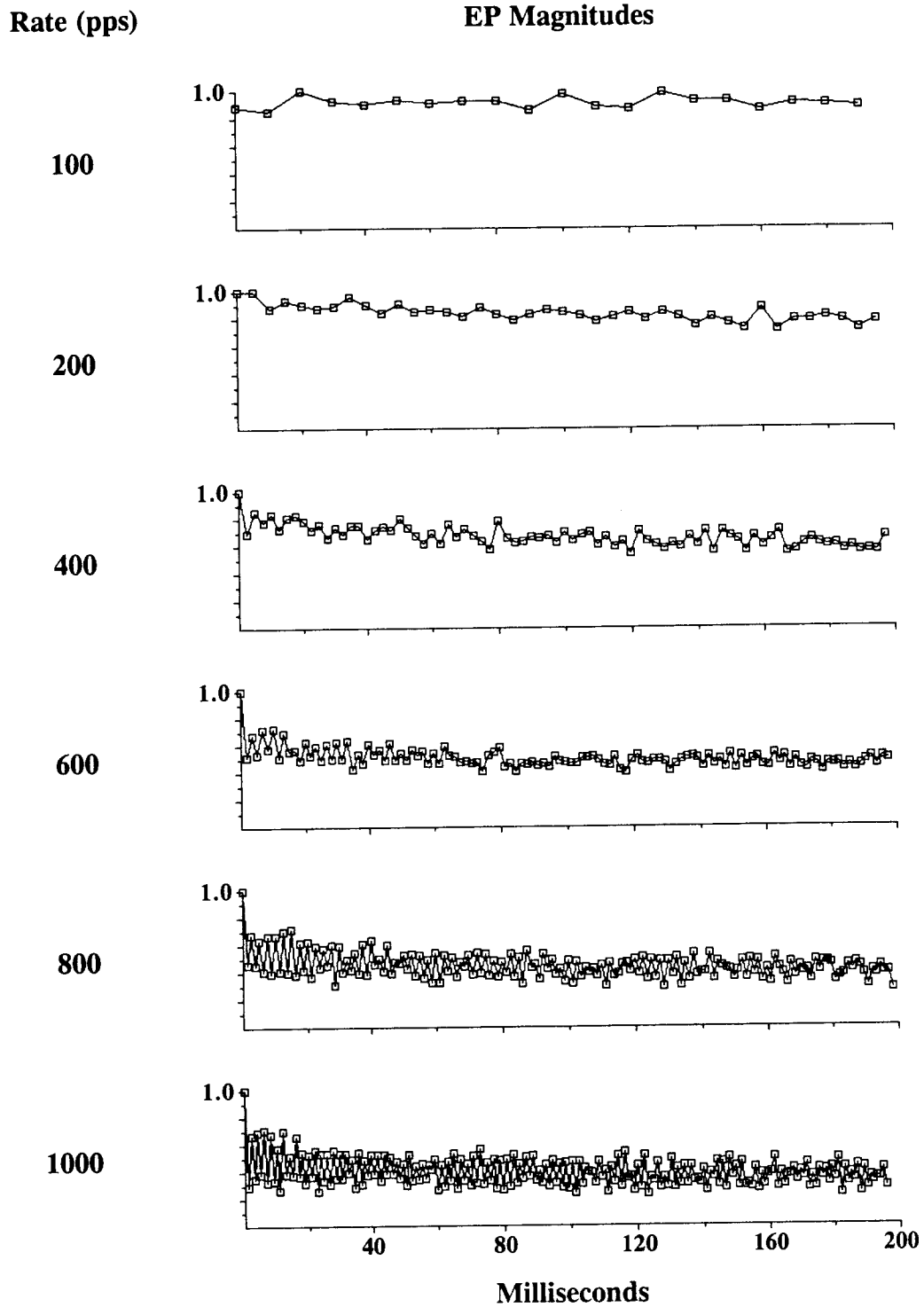

Fig. 11. Magnitudes of EPs derived from the recordings of Fig. 9. Magnitudes for the entire 200 ms duration of each record are shown.

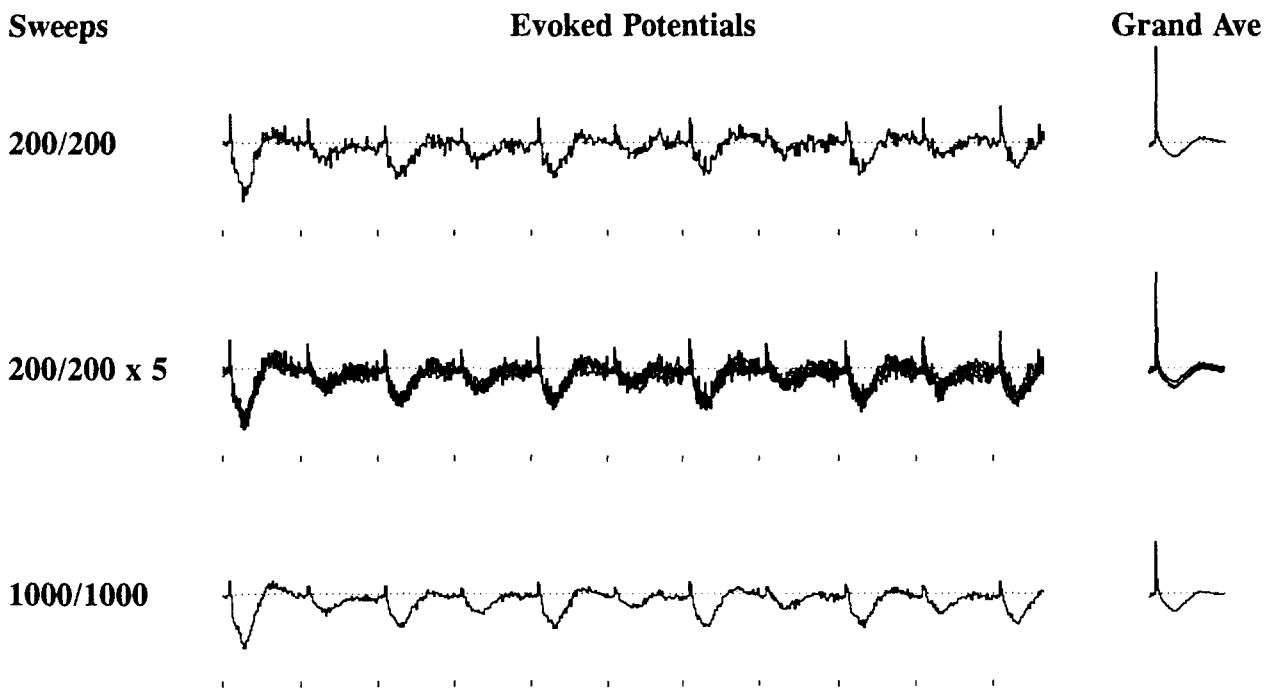

Fig. 12. Repeatability of EP measures. The top panel reproduces the record shown in Fig. 9 for the pulse rate of 1000/s. The middle panel shows five repeated measures (most on different days) with 200 sweeps for each pulse polarity. The bottom panel is a single record combining all available data for this stimulus condition. Note that the recordings are highly repeatable (middle panel) and that the noise level in the recordings can be reduced substantially by increasing the number of sweeps (bottom panel).

alternating response approaches the apparent $\pm 10\%$ variation in EP magnitudes for the 200-sweep recordings, that variation masks visualization of the alternating pattern.

The inability of the nerve to reflect fully the stimulus waveform for rates at and above 400 pps may have a psychophysical correlate. Fig. 14 shows results from a magnitude estimation experiment, in which the same subject (SR2) was asked to nominate a pitch for each stimulus in a set of stimuli with different pulse rates. The subject was instructed to nominate a low number for a low pitch and a high number for a high pitch. The stimuli included 200 ms bursts of pulses presented (as above) to electrode 3 at the indicated rates. The different rates were presented in a randomized order according to the method of constant stimuli. The number of trials for each stimulus condition was 30. Note that the mean of pitch judgments doubles with a doubling of rate from 100 to 200 pps, and from 200 to 400 pps. Further increases in rate do not produce similar increases in adjudged pitch. Increases in rate beyond 1600 pps do not produce statistically different judgments in pitch. Such saturation of pitch judgments has been found in many other studies with cochlear implant patients (e.g., Shannon, 1983) and is referred to as the "pitch saturation limit." The psychophysical data of Fig. 14 indicate a

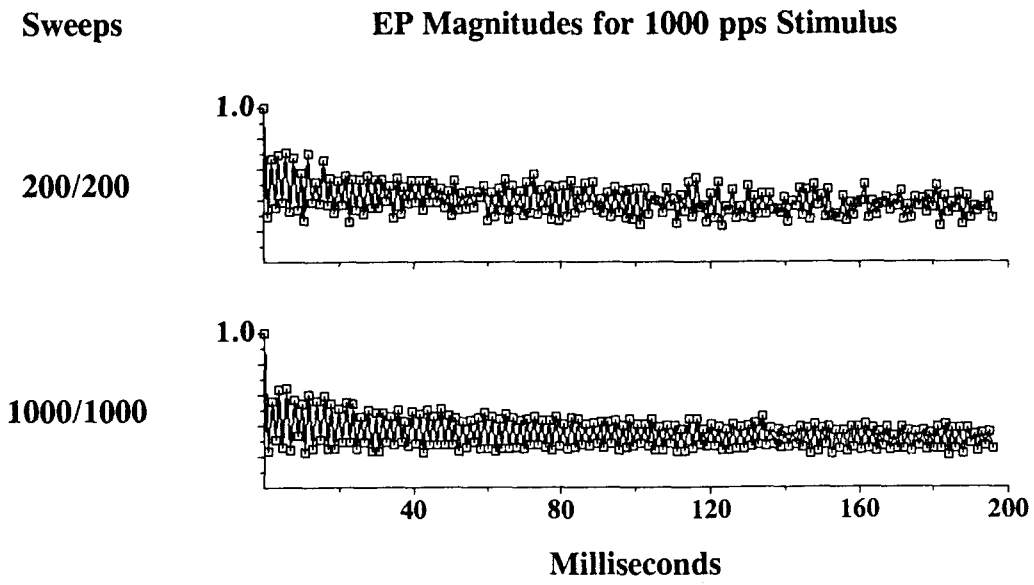

Fig. 13. Magnitudes of EPs derived from recordings obtained with 200 or 1000 sweeps for each polarity of stimulus pulses. The 200 sweeps recording is presented in the bottom panel of Fig. 9 and the top panel of Fig. 12, and the 1000 sweeps recording is presented in the bottom panel of Fig. 12. Pulse rate was 1000/s.

compression of pitch judgments between 400 and 800 pps. The EP data of Figs. 9 and 11 demonstrate a failure of the auditory nerve to follow pulses with equal magnitudes of response for rates of stimulation higher than 400 pps. An alternating pattern of response is clearly evident in the initial parts of the records for the 600 and 800 pps conditions (with 200 sweeps, Fig. 11), and for the full 200 ms of the record for the 1000 pps condition (with 1000 sweeps, Fig. 13). The pitch saturation limit may be related to such patterns of response at the auditory nerve.

Additional experiments with subject SR2 have provided further evidence of a possible connection between the peripheral representation and pitch perception. EPs also have been measured with stimulation of his electrode 1 (the apicalmost electrode in the Ineraid array). Those EPs show a better ability of the local neural population to follow pulses presented at relatively high rates, i.e., the onset of the alternating response occurs at about 600 pps with stimulation of electrode 1, whereas that onset is at about 400 pps for stimulation of electrode 3, as described above. A magnitude estimation experiment was conducted in which SR2 was asked to judge pitch for stimuli presented at different rates to either electrode 1 or electrode 3. As might be expected, a higher limit of pitch saturation was found for electrode 1. At low rates pitch judgments were lower for electrode 1, consistent with its more apical placement, and at high rates (above 400 pps) pitch judgments were higher for electrode 1, consistent with the better ability of the neural population stimulated by electrode 1 to follow the pulses at those rates.

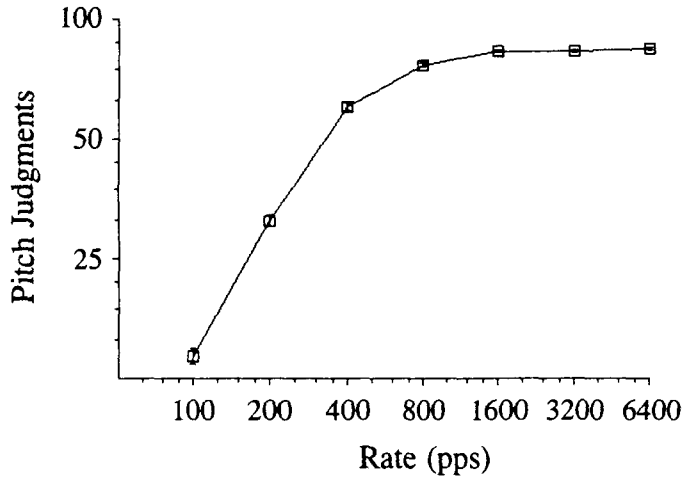

Fig. 14. Results from a magnitude estimation experiment, in which the subject (Ineraid subject SR2) was asked to nominate a pitch for each stimulus. The subject was instructed to nominate a low number for a low pitch and a high number for a high pitch. Stimuli included 33 μs/phase pulses presented to electrode 3 at the indicated rates within a 200 ms burst of pulses. The pulse amplitude for each rate was adjusted to produce a most comfortable loudness (MCL) percept. Once the amplitudes for MCL percepts had been identified, the stimuli for the various rates were played in sequence. Fine adjustments in pulse amplitudes were then made, and the sequence repeated, until all stimuli were judged to be equally loud. In the experiment, the stimuli were presented in a randomized order according to the method of constant stimuli. The number of trials for each stimulus condition was 30. The bars (within the symbols for most conditions) show standard errors of the mean for the judgments.

Such conjecture about a possible connection between the peripheral representation and pitch judgments must be tempered by the fact that the stimuli used in the EP and psychophysical experiments were not identical. EP measures were made with one amplitude of pulses across rates, whereas the amplitudes used in the psychophysical experiments were somewhat different across rates, to achieve a balance of loudnesses. Also, the duration of pulses for the EP measures was 16.4 μs/phase, whereas the duration for the psychophysical experiments was 33 μs/phase. A direct comparison between the psychophysical and EP results would require the use of identical stimuli. We have included such direct comparisons in our plans for future studies.

We note that results from EP measurements with Ineraid subject SR3 were quite similar to the results presented above for subject SR2. EPs in response to stimulation of her electrode 1 also indicated a better ability of the local population to follow pulses at relatively high rates, compared with EPs in response to stimulation of electrode 3. Waveforms and magnitudes of the EPs were similar for the two subjects as well.

In contrast, EPs recorded in studies with Ineraid subject SR10 showed a relatively poor ability of the nerve to follow pulses much above 200 pps, much lower amplitudes of EPs (about 1/5 the amplitudes found for SR2 and SR3), and a somewhat broader waveform of the EPs compared with the waveforms for the other two subjects. Also, the alternating pattern of responses seen at 800 and 1000 pps for subjects SR2 and SR3 was supplanted in the results for SR10 by more complex patterns, similar to those predicted in the model results of Fig. 5. We note that SR10 has relatively high thresholds, relatively narrow dynamic ranges between threshold and most comfortable loudness (MCL) percepts, and relatively poor speech recognition performance with his implant compared with those measures for subjects SR2 and SR3. The ratios of EP magnitudes for the first two pulses in the stimuli at the different rates for subject SR10 indicate slower recovery from prior stimulation than is the case for SR2 and SR3. The lower amplitudes of the EPs, along with a broadening of the EP waveform, suggest the possibility of a less synchronized response among stimulated neurons for SR10. The different pattern of EP magnitudes across pulse rates observed for SR10 also suggests that his population of neurons has properties that are different from those of SR2 and SR3. Further studies with additional subjects will be required to evaluate possible correlations between speech reception scores and recorded patterns of intracochlear EPs, and between psychophysical measures and EP patterns.

Responses to SAM Pulse Trains

Recordings of EPs in response to SAM pulse trains are presented in Figs. 15 and 16, for the indicated modulation frequencies and carrier rates. These figures show results for subject SR2, with stimulation of his electrode 3 (as in Fig. 9). Results for subject SR3 were quite similar to those shown here for SR2. The data for SR10 are still being plotted and analyzed.

The patterns of EPs in Figs. 15 and 16 are similar, in some cases quite similar, to the predicted patterns in Figs. 6 and 7. These similarities are further illustrated in Figs. 17 to 20, which show direct comparisons of EP magnitudes with model predictions. The standard conditions were used for the model simulations, with the same ratios of peak pulse amplitudes used in the experiments of Figs. 15 and 16. Data and model predictions are compared for the first 50 ms of the stimuli in Figs. 17-20. Pearson correlations of data and model predictions are presented to the right for the first 50 ms (r_{50}) and the entire 200 ms (r_{200}) for each condition. Note the general agreement between model and data, as indicated by correlation coefficients of 0.88 or higher. In some cases, the EPs show a greater response to a second, high amplitude pulse than predicted by the model. For the carrier rate of 500 pps and the modulation frequency of 100 Hz, for instance, the model predicts a large response for the third pulse in the stimulus and a greatly diminished response to the fourth pulse, a pulse with the same amplitude as the third pulse. The EP data show a much smaller reduction in the response to the fourth pulse. As with trains of constant amplitude pulses (Fig. 9), agreement between model and data might be improved by different choices of parameters for the model, such as the different values of absolute and relative refraction derived from the data of Fig. 10, or by the addition of a low level of membrane noise.

We note that percepts of SAM pulse trains have been highly consistent with the recorded and predicted patterns of neural response. In magnitude estimation experiments, SR2 judges 250 pps pulse trains modulated with 200 Hz sinusoids as being statistically identical in pitch to 250 pps pulse trains modulated with 50 Hz sinusoids. He also judges 250 pps pulse trains modulated with 150 Hz sinusoids

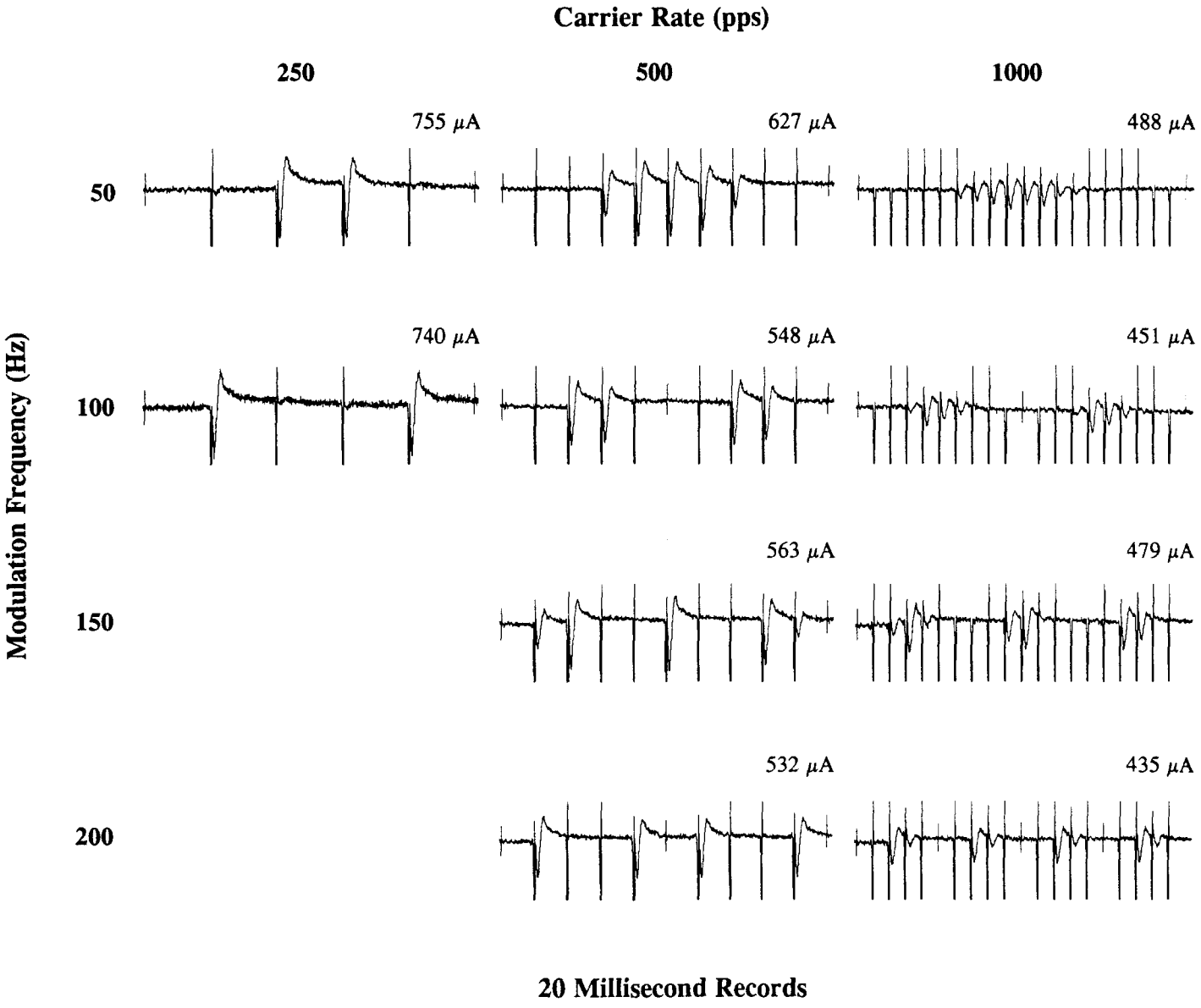

Fig. 15. Recordings of intracochlear evoked potentials for Ineraid subject SR2. Stimuli included sinusoidally amplitude modulated (SAM) pulse trains with the indicated modulation frequencies and carrier rates. The duration of pulses in the carriers was 32.8 μs/phase, and the depth of modulation 100 percent. Stimuli were delivered to electrode 3 in the Ineraid implant, and potentials were recorded differentially between electrode 4 and the ipsilateral mastoid. The amplitude of the carrier pulses was adjusted for each condition to produce a most comfortable loudness (MCL) percept. The amplitudes for each of the conditions are indicated above the traces. Responses from 400 sweeps of the stimulus with the positive phase leading for each pulse were added to responses from 400 sweeps of the stimulus with the negative phase leading. No blanking circuit was used. Timing of pulse presentations is indicated by vertical lines. The full amplitudes of many of these pulse artifacts are not shown in the figure (i.e., the artifacts are clipped in the figure). Note the general agreement between the magnitudes of the EPs shown in this figure and the number of spikes predicted for each pulse in Fig. 6.

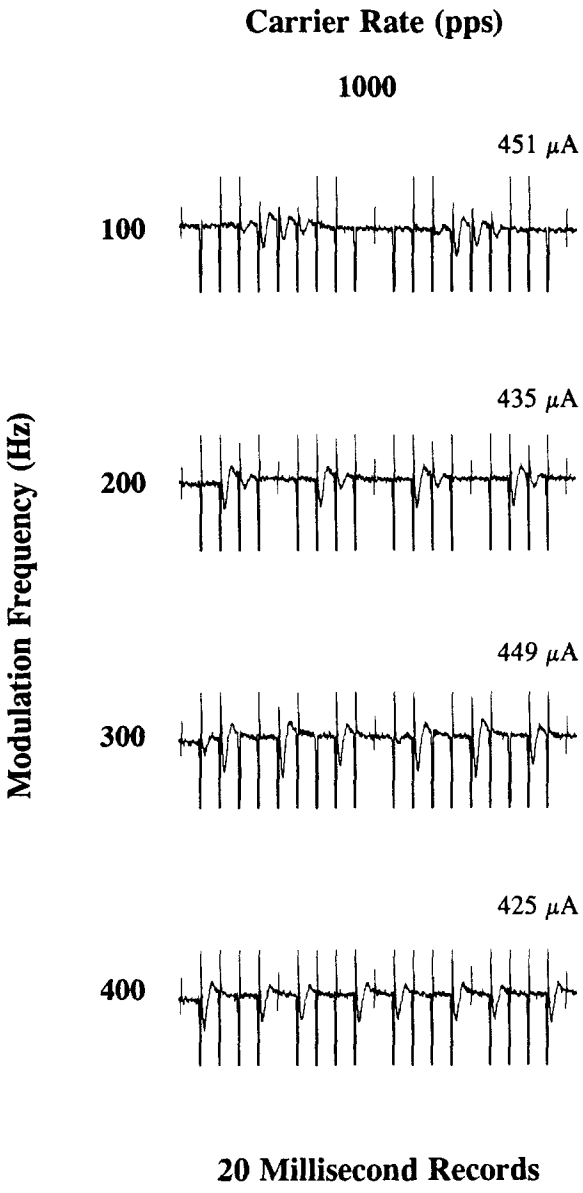

Fig. 16. Recordings of intracochlear evoked potentials for Ineraid subject SR2. Stimuli included sinusoidally amplitude modulated (SAM) pulse trains with the indicated modulation frequencies and carrier rates. Except for the two additional modulation frequencies, conditions of stimulation and recording are the same as those in Fig. 15. Amplitudes of the carrier pulses for each of the conditions are indicated above the traces. Note the general agreement between the magnitudes of the EPs shown in this figure and the number of spikes predicted for each pulse in Fig. 7.

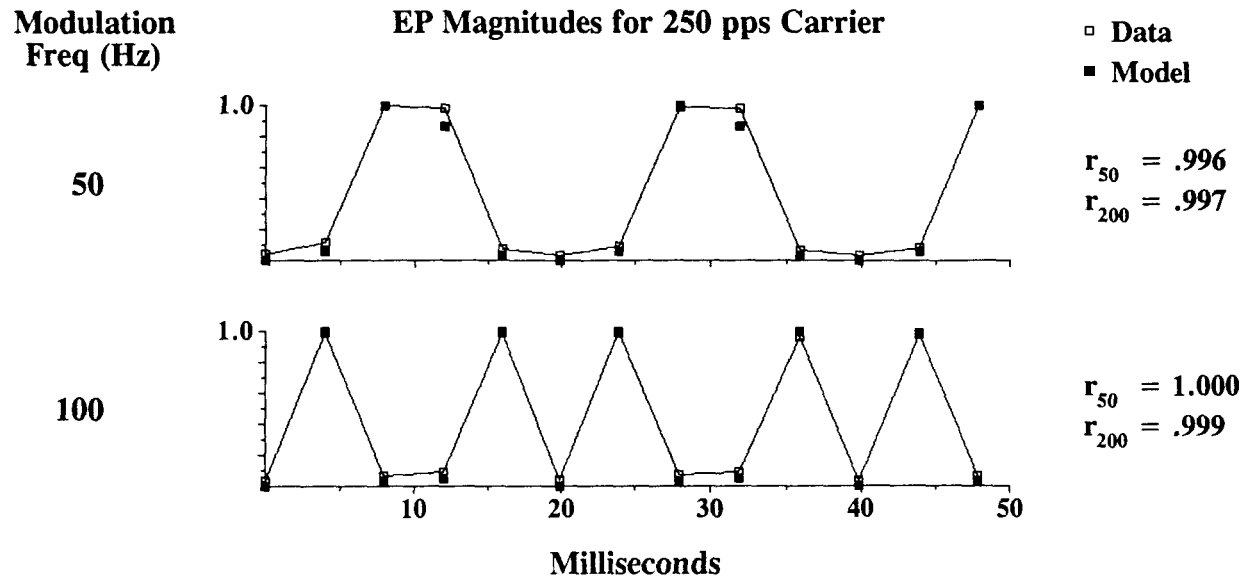

Fig. 17. Comparison of evoked potential magnitudes with model predictions of the number of neurons responding to each pulse in sinusoidally amplitude modulated (SAM) trains of pulses. The EP magnitudes were derived from the recordings presented in Fig. 15, for the carrier rate of 250 pps. The standard conditions were used for the model simulations (see text), with the same ratios of pulse amplitudes used in the experiment of Fig. 15. Data and model are compared for the first 50 ms of the stimuli in this illustration. Pearson correlations of data and model are presented to the right for the first 50 ms (r_{50}) and for the entire 200 ms (r_{200}) for each condition.

to be identical in pitch to 250 pps pulse trains modulated with 100 Hz sinusoids. These results are consistent with the predicted effects of aliasing for these conditions, as shown in Fig. 6. For 500 pps carriers, SR2 scales pitch monotonically with increases in modulation frequency. However, SR2 describes the percept for the 150 Hz modulation condition as sounding "very rough and complex." Also, the percept for the 200 Hz modulation condition is described as having "at least two separate tones." These observations are consistent with the complex pattern of responses predicted for the 150 Hz modulation condition, and with the two distinct intervals between large responses predicted for the 200 Hz modulation condition. For 1000 pps carriers, the percepts for the 150 and 200 Hz modulation conditions are described as relatively "smooth and tonal," particularly for the 200 Hz modulation condition. These descriptions are consistent with the more uniform distribution of predicted and recorded neural responses across peaks in the modulation waveforms for these stimuli (see Figs. 6 and 15). For modulation frequencies of 300 and 400 Hz, however, a "rough and complex" percept is again reported for the 300 Hz modulation condition and a multitonal percept is again reported for the 400 Hz modulation condition. Figs. 7 and 16 show a complex pattern of neural responses for the 300 Hz modulation condition and a pattern with two distinct intervals between large responses for the 400 Hz modulation condition.

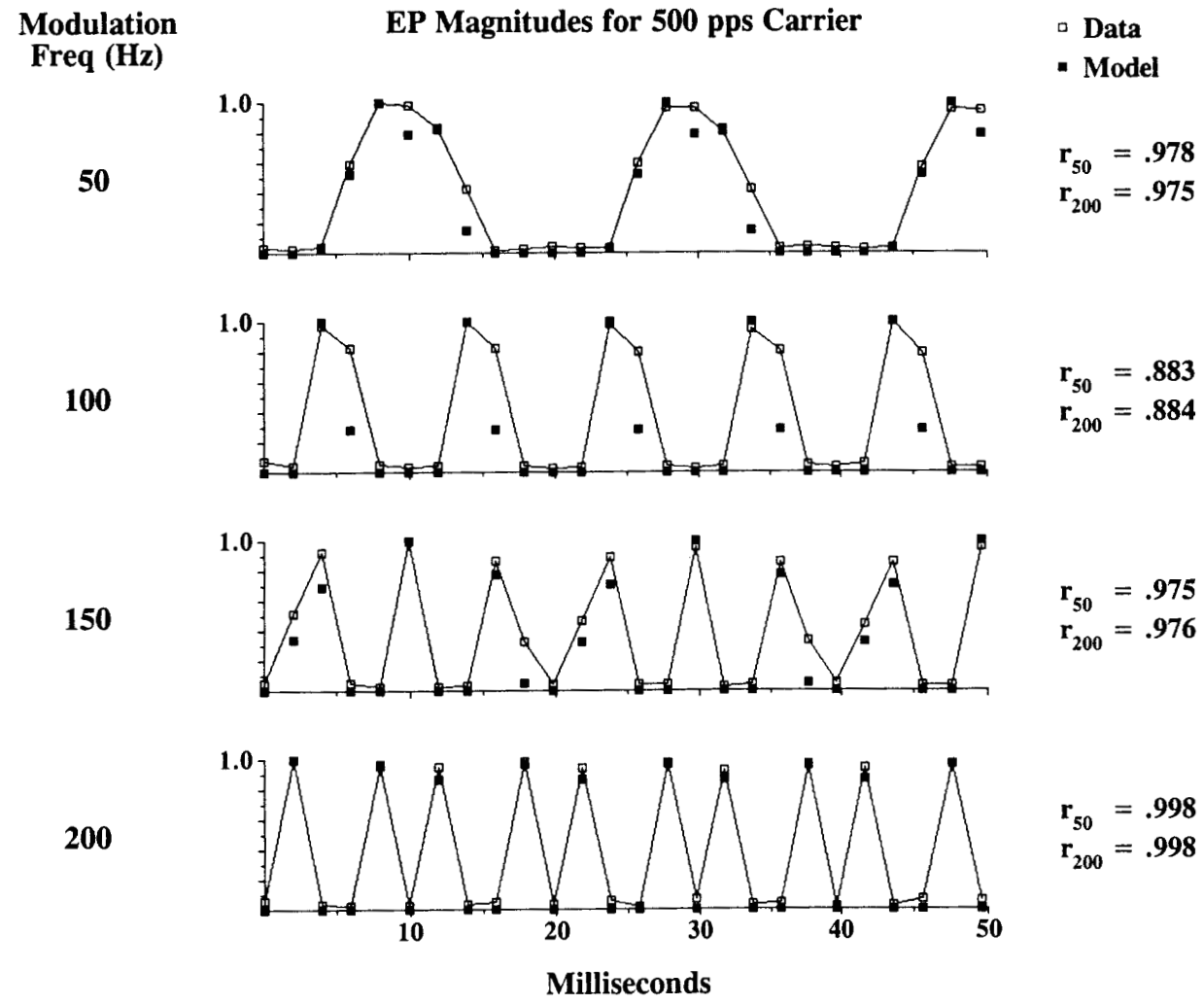

Fig. 18. Comparison of evoked potential magnitudes with model predictions of the number of neurons responding to each pulse in sinusoidally amplitude modulated (SAM) trains of pulses. The EP magnitudes were derived from the recordings presented in Fig. 15, for the carrier rate of 500 pps. Parameters for model simulations and organization of the panels are the same as those in Fig. 17.

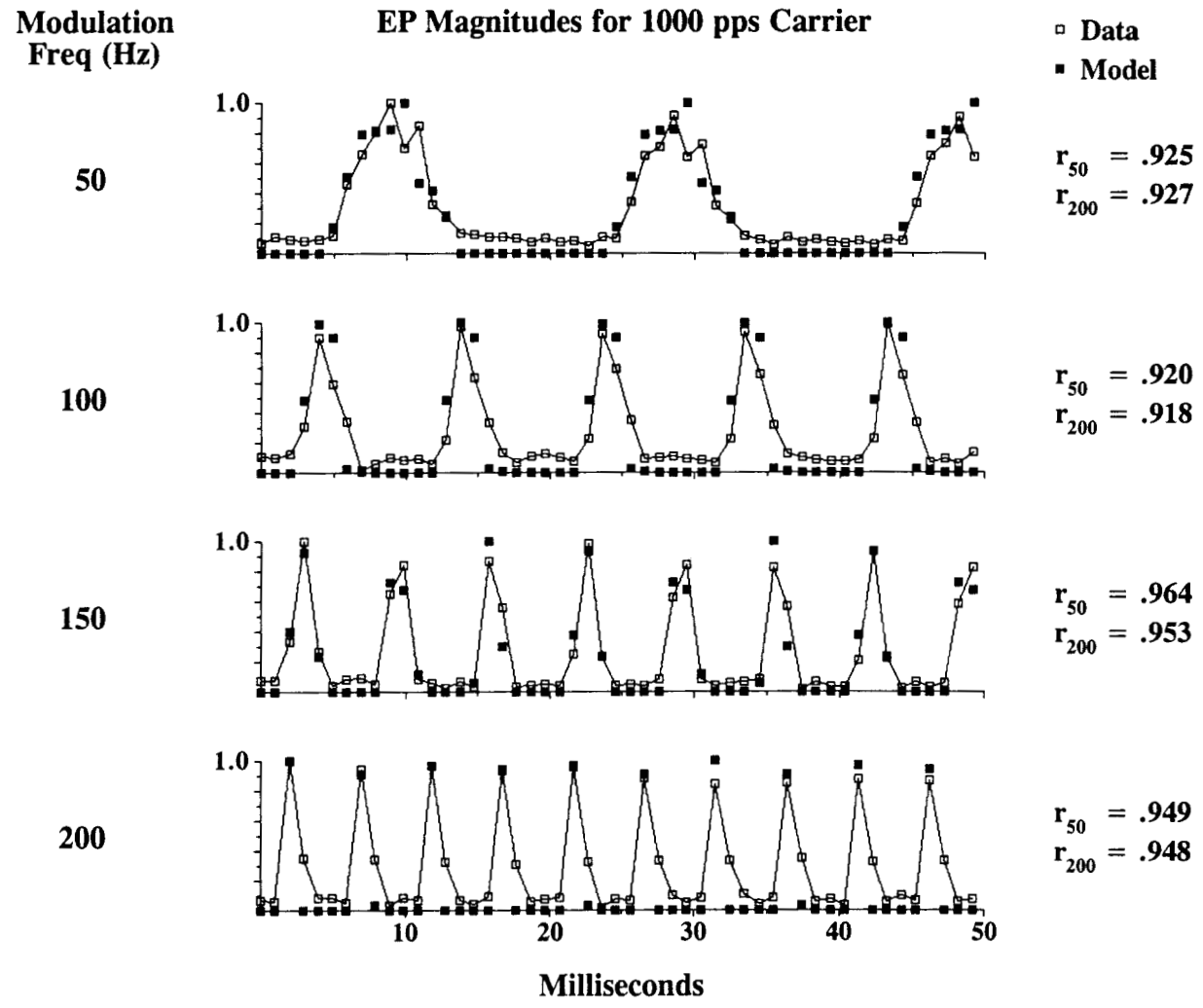

Fig. 19. Comparison of evoked potential magnitudes with model predictions of the number of neurons responding to each pulse in sinusoidally amplitude modulated (SAM) trains of pulses. The EP magnitudes were derived from the recordings presented in Fig. 15, for the carrier rate of 1000 pps. Parameters for model simulations and organization of the panels are the same as those in Fig. 17.

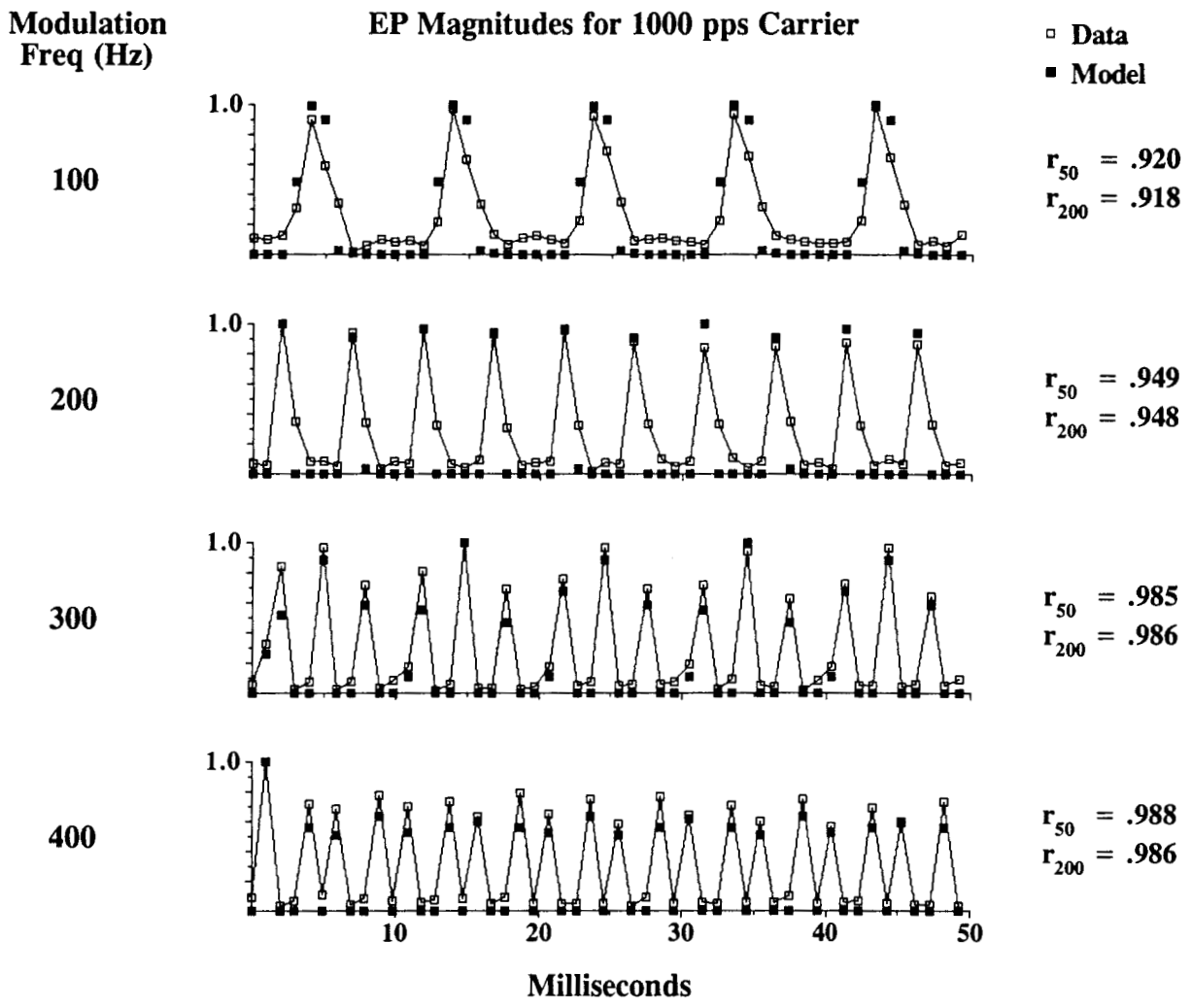

Fig. 20. Comparison of evoked potential magnitudes with model predictions of the number of neurons responding to each pulse in sinusoidally amplitude modulated (SAM) trains of pulses. The EP magnitudes were derived from the recordings presented in Fig. 16, for the carrier rate of 1000 pps. Parameters for model simulations and organization of the panels are the same as those in Fig. 17.

Further Comparisons of Model Predictions and Data

The availability of a direct measure of whole-nerve responses in the electrically stimulated cochleas of human subjects, for a wide variety of stimuli, makes possible a rapid convergence of model predictions of those responses with data. In the presentations of EP results above, we have provided initial comparisons between model predictions and data for the standard conditions of the model. The observed differences between model and data have motivated additional modeling studies, using (a) a model with the new recovery function for single neurons suggested in the caption and accompanying text for Fig. 10 and (b) models with the standard conditions along with the addition of various levels of membrane noise, as in Fig. 8. In addition, each single parameter in the model has been varied to determine its effect on predicted patterns of response to trains of pulses.

Pulse Trains

Predicted patterns of response to pulse trains are presented in Figs. 21 and 23 for the new recovery function and for various levels of membrane noise, respectively. The conditions of stimulation correspond to those in Figs. 9 and 11. The model simulations include four different amplitudes of stimulus pulses, as the prior results in Fig. 5 indicated a possible sensitivity to pulse amplitude. Note that the amplitude of best fit for the model may not correspond to the actual amplitude of pulses used in the EP experiments because the amplitudes in the model are somewhat arbitrary in the sense that the amplitude required to produce a given number of spikes depends on a variety of parameters (e.g., a reduction in chronaxie can reduce dramatically that amplitude for short-duration pulses). The lowest amplitude in the simulations corresponds to a level required to produce approximately 2300 spikes in the first 10 ms of the record for the 1000 pps condition. The difference between this level and the level required to produce any spikes generally is between 3 and 5 dB for simulations without the addition of membrane noise and more in the presence of such noise. Each successively higher amplitude is 1.5 times greater than the one that precedes it. The total dynamic range of the amplitudes is 10.6 dB. This, along with the difference between the lowest level and the level at which spikes are just elicited, corresponds to a dynamic range from a presumed threshold to loud percepts of 13 to 16 dB, consistent with the dynamic ranges of some of the better implant patients for short pulses. As in prior figures, results are normalized to the greatest number of spike counts per pulse in each 200 ms record.

In general, substitution of the new recovery function (period of absolute refraction changed from 700 to 488 μs and time constant of relative refraction changed from 1250 to 1548 μs) produced only subtle effects in the predicted patterns of response. For example, the depth of the alternating response for the 200 and 400 pps conditions is somewhat greater with the new recovery function. Use of the new recovery function does not improve the fit of model to data. This lack of improvement is illustrated further in Fig. 22, which shows in the left column the EP magnitude data (from Fig. 11), in the middle column model predictions for the standard conditions and a pulse amplitude of 1125 μA (corresponding to a level above threshold appropriate to a most comfortable loudness at 1000 pps), and in the right column model predictions for the new recovery function and the same pulse amplitude. Note that few of the features in the data are predicted by either variation of the model for pulse rates at and above 400/s.

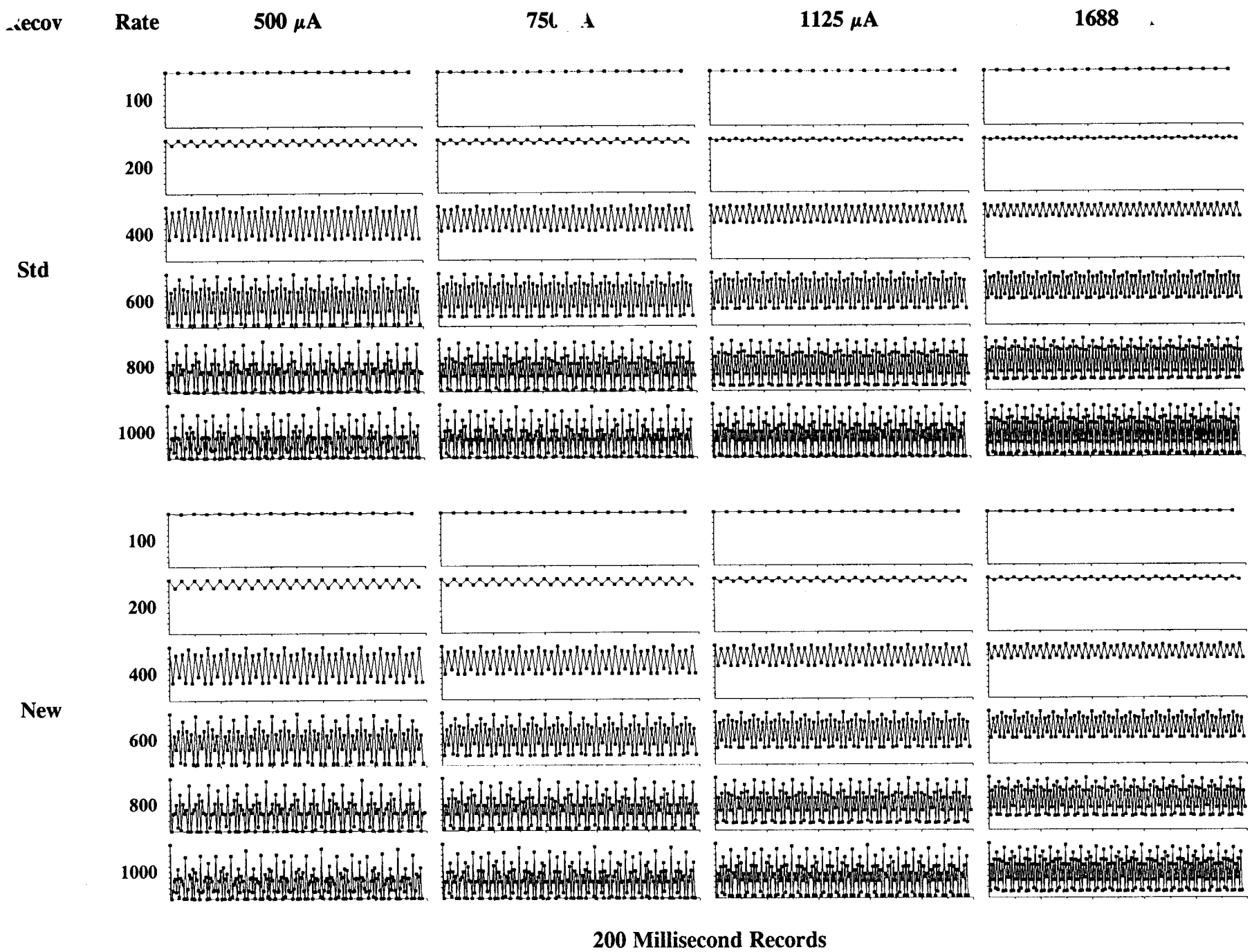

Fig. 21. (please see caption on next page)

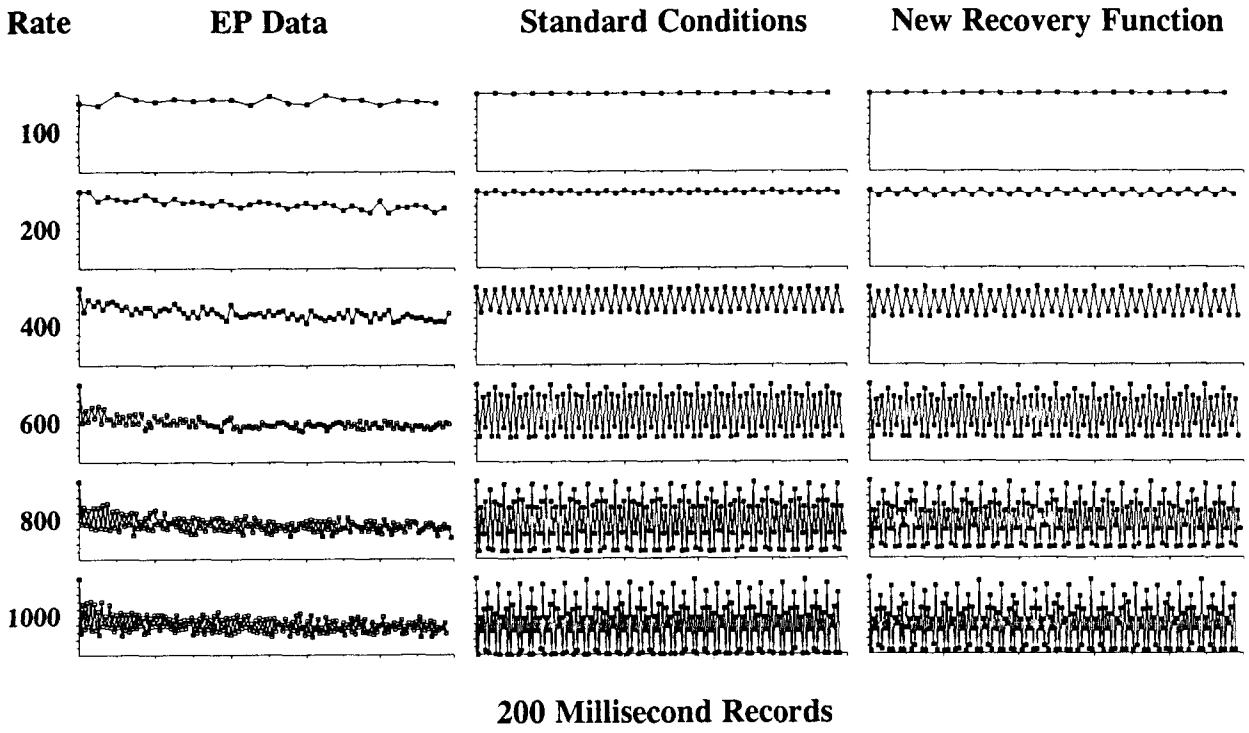

Fig. 22. Comparison of EP magnitude data (from Fig. 11) with model predictions. Predictions from the model using the standard conditions are presented in the middle column and predictions from a modified model using the new recovery function for each neuron are presented in the right column. Pulse amplitude for the model simulations was 1125 μA.

Fig. 21. (from the prior page) Predicted magnitudes of neural responses for each pulse in trains of pulses. The upper panels show predictions from a model using the standard conditions, for the indicated pulse rates and amplitudes. The lower panels show predictions from a modified model with different parameters for the recovery function of single neurons (period of absolute refraction changed from 700 to 488 μs and time constant of relative refraction changed from 1250 to 1548 μs). Magnitudes are normalized to the maximum response for each panel.

As might be expected from Fig. 8 and the accompanying discussion, addition of even small amounts of membrane noise in the model can produce large effects on predicted patterns of response. Indeed, a remarkable improvement in the fit between model and data is observed with the addition of membrane noise. The region of best fit appears to be in the neighborhood of the panel in Fig. 23 showing responses predicted for addition of membrane noise at the "2.0" level (corresponding to that level in Fig. 8) and for the stimulus amplitude of 1125 μA (Fig. 23, part 3, third column).

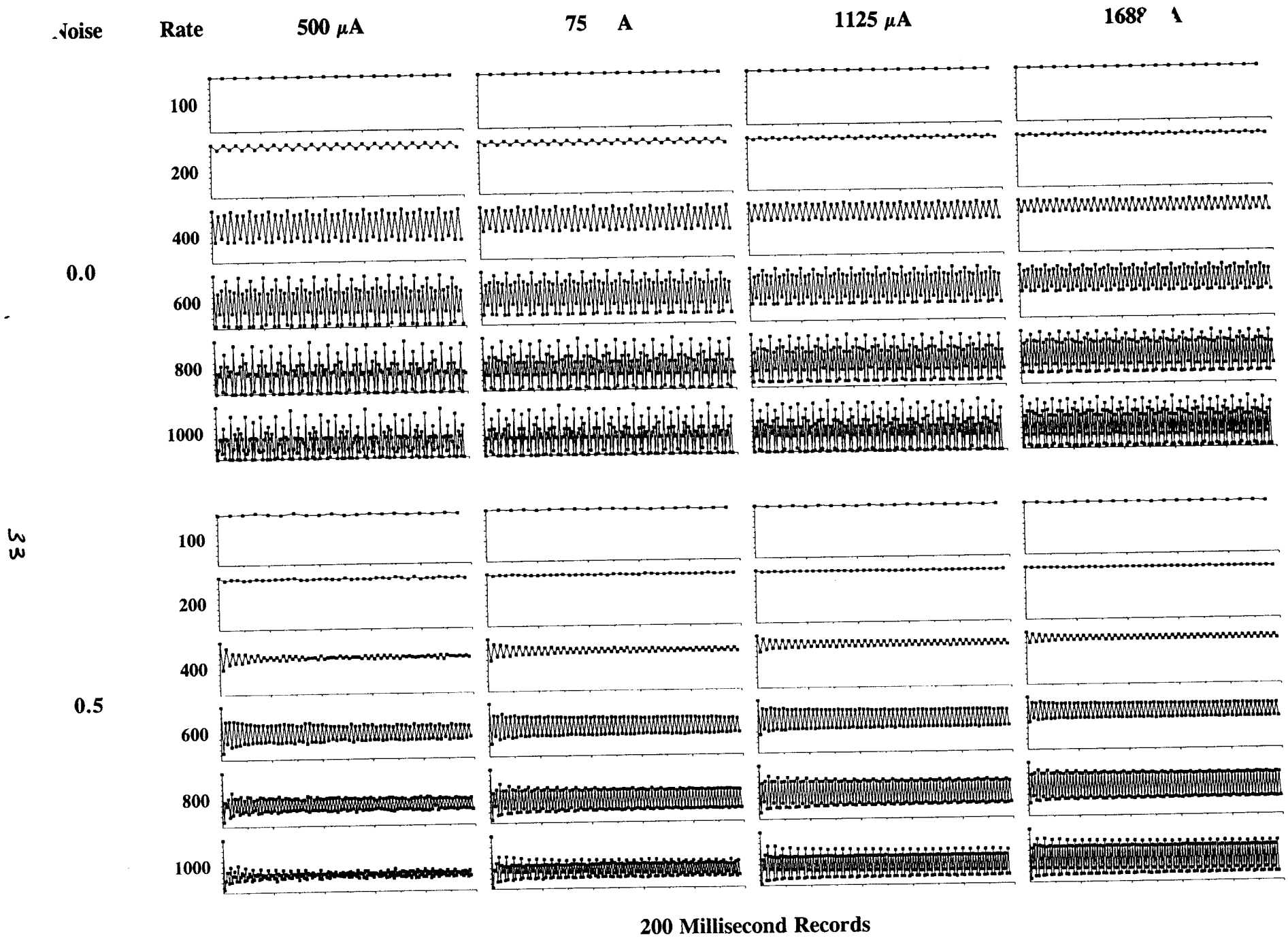
Noise
Rate
500 μA
75 A
1125 μA
1688 A
0.0
0.5
100
200
400
600
800
1000
200 Millisecond Records

Fig. 23. (part 1)

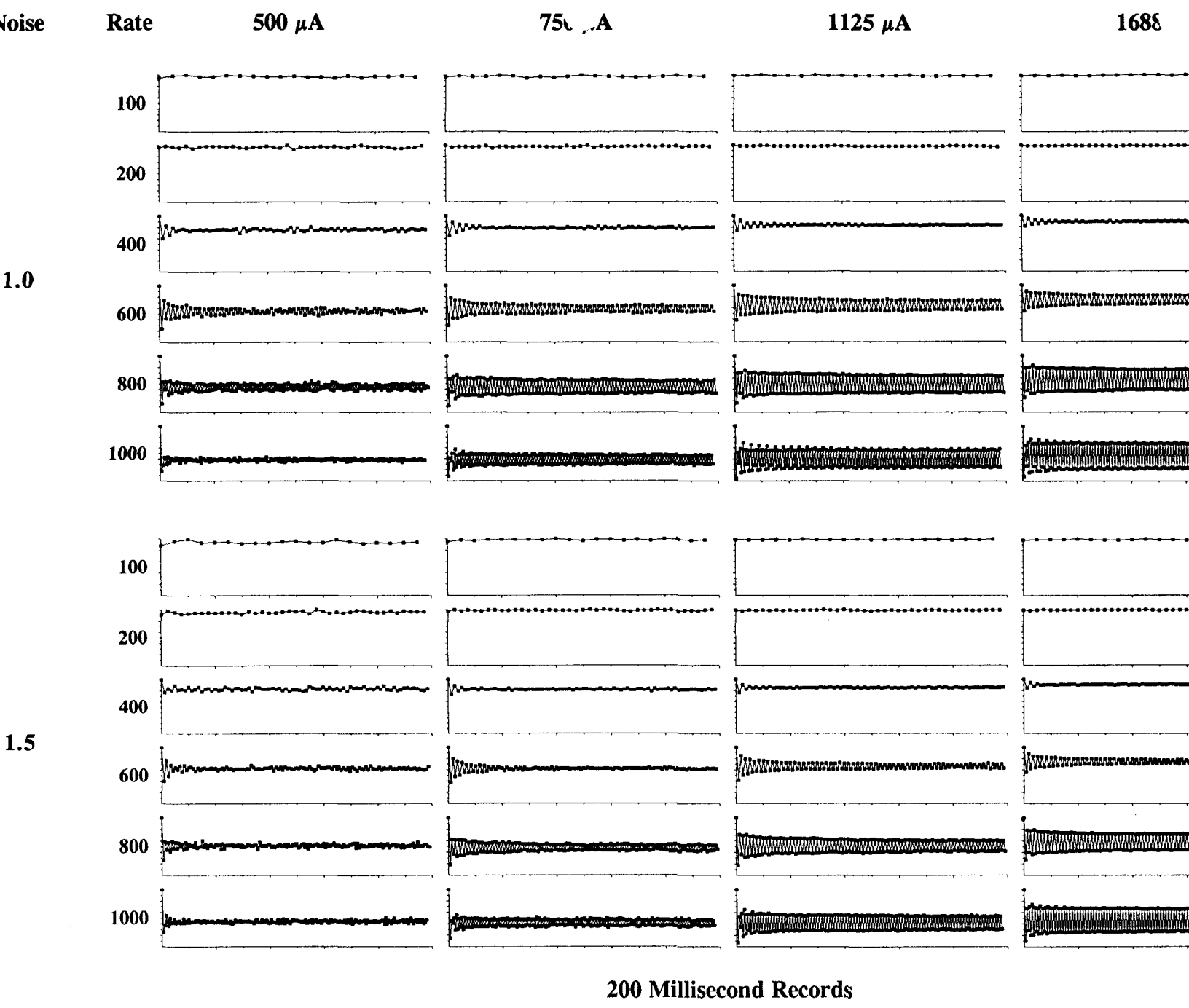
Noise
Rate
500 μA
1125 μA
1.0
100
200
400
600
800
1000
1.5
100
200
400
600
800
1000
200 Millisecond Records

Fig. 23. (part 2)

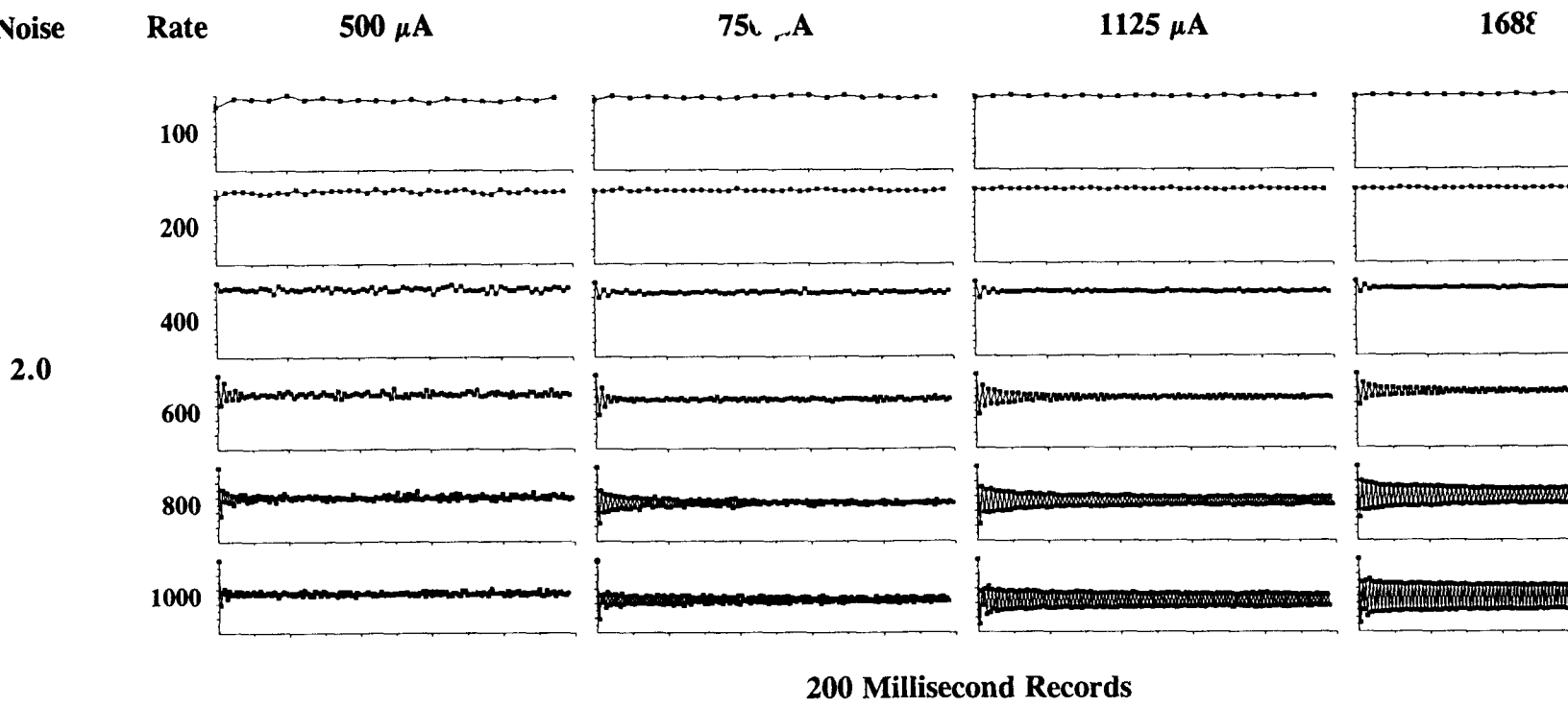

Fig. 23. (part 3)

Fig. 23. Predicted magnitudes of neural responses for each pulse in trains of pulses. The upper panels in part 1 of the figure sho from a model using the standard conditions, for the indicated pulse rates and amplitudes. The remaining panels show predictions using the standard conditions along with the addition of various levels of membrane noise, as indicated to the left of each s Magnitudes are normalized to the maximum response for each panel.

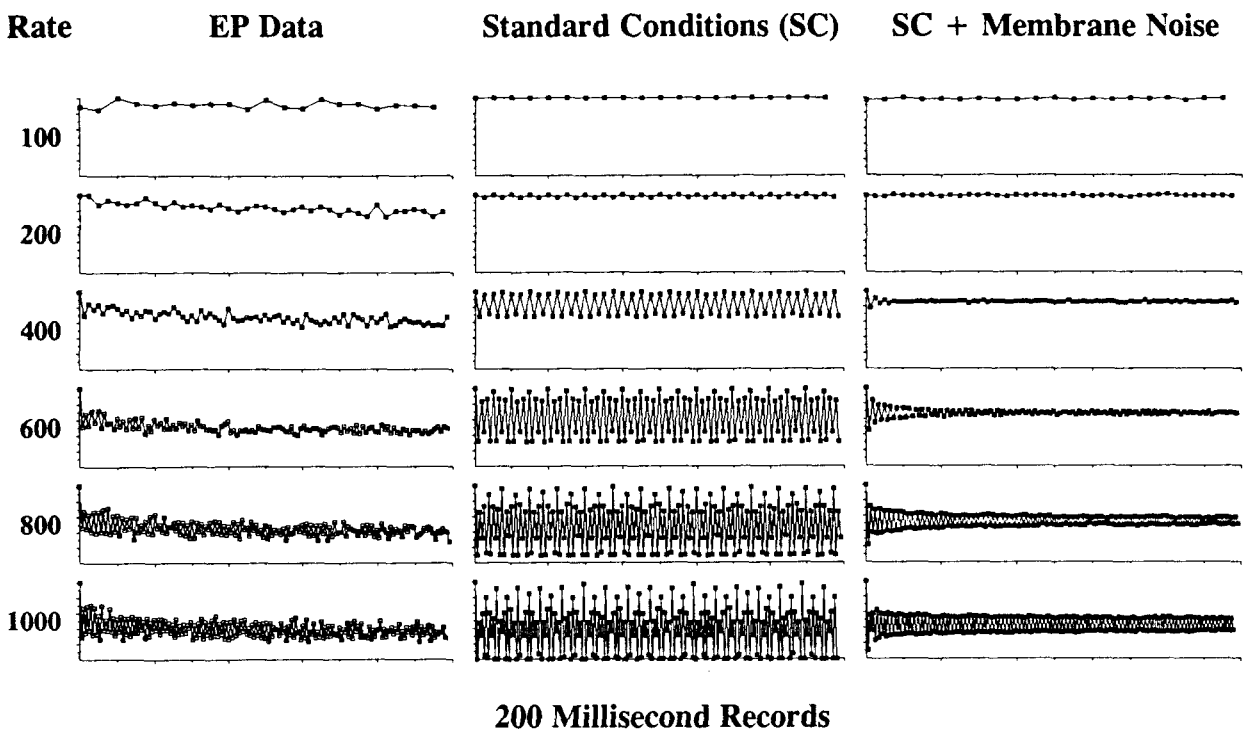

Fig. 24. Comparison of EP magnitude data (from Fig. 11) with model predictions. Predictions from the model using the standard conditions are presented in the middle column and predictions from a model using the standard conditions plus membrane noise are presented in the right column. The level of membrane noise was 2.0. The pulse amplitude used for both sets of model simulations was 1125 μA.

The agreement between model and data is illustrated further in Fig. 24, which shows EP magnitude data in the left column, model predictions using the standard conditions in the middle column, and model predictions with the standard conditions plus membrane noise in the right column. The stimulus amplitude for the two models was 1125 μA and the noise level for the "SC + Membrane Noise" model was 2.0. For the 100 and 200 pps conditions, the model with membrane noise predicts a relative uniformity of responses across pulses. Some stochastic variation in responses from pulse to pulse also is predicted. At intermediate rates, an initial decrement in the response is predicted for the second pulse, followed by an alternating pattern of response for some period thereafter. This period is brief for stimulation at 400 pps, lasting only several cycles into the burst before it is damped. The period is lengthened for stimulation at 600 pps and the magnitude of the initial alternations is increased. The alternating response disappears about 40 ms into the burst. At 800 pps, a large alternating response is predicted for the first 40 ms, followed by a relatively uniform pattern of alternating responses. The amplitude of the alternations in the 40-200 ms interval is less than the amplitude of the alternations in the first 40 ms. At 1000 pps, a large alternating response is predicted for the entire 200 ms of the burst. The predicted pattern of response is relatively uniform after the sixth or seventh pulse, with only a slight dampening of the alternations with increasing time into the burst.

These and other features of predicted responses reflect features observed in the EP data. In particular, the recorded responses show a maintenance of EP magnitudes across pulses for the 100 and 200 pps conditions. At 400 pps, the data show an initial decrement in the response following pulse 1, with a possible oscillation in the response for several subsequent pulses (this impression needs to be verified with recordings using a higher number of sweeps, as the noise variation in the present recordings is comparable to the level of variation in the predicted pattern of alternating responses for the 400 pps condition). At 600 pps, the data show an initial alternation in the response followed by uniform responses after approximately 40 ms. A still larger alternation in response is observed for the first 40 ms for the 800 pps condition. Variations in response magnitudes thereafter are larger than the variations observed for the same period in the 600 pps record. This increased variation suggests the possibility of a continuing pattern of alternating responses for the 800 pps condition, although a strict alternation in responses during this interval is not demonstrated by the present data, collected with 200 sweeps for each pulse polarity. Noise in the recordings for this relatively small number of sweeps may have masked or distorted an alternating pattern of response. At 1000 pps, a quite large alternation in the response is seen for the first 40 ms. The variation in response magnitudes diminishes somewhat over the remainder of the record but remains large.

A feature of the data not predicted by the model (with or without membrane noise) is the gradual decrement in the response over time for rates of stimulation at and above 200 pps. As mentioned before, this decrement probably reflects long-term accommodation in the response at those rates. Inclusion of a model element to account for accommodation, such as the simple element suggested by Hill [1936], might improve the fit between model and data for this feature.

Additional features of the data and model predictions are illustrated in Figs. 25 to 27. Fig. 25 shows data and model predictions for stimulation at 1000 pps, with the data from the combined records for that condition. As indicated in Fig. 13, the total number of sweeps was 1000 for each pulse polarity in the combined records. In contrast to the 200 sweeps record of Fig. 24, the 1000 sweeps record in Fig. 25 (left column) shows an alternating pattern for the entire 200 ms of the burst, in agreement with model predictions (Fig. 25, right column). The improvement in signal-to-noise ratio provided with the higher number of sweeps allows demonstration of subtle features in the response.

Fig. 26 shows data and model predictions for the first 50 ms of the burst for each stimulus condition. The data records are those obtained with 200 sweeps for each pulse polarity, as in Fig. 24. These expanded records show in detail the initial alterations in EP magnitudes for the 600, 800 and 1000 pps conditions (left column). Such alterations also appear in the predictions of the model using the standard conditions plus membrane noise (right column).

Agreement between model predictions and data is illustrated further in Fig. 27, where the data are from the first 50 ms of the 1000 sweeps record for the 1000 pps condition. The alternating pattern of response in the data is quite similar to the alternating pattern in the model predictions. Both data and predictions show an initial decrement in the response at pulse 2, followed by the alternating response. The amplitude of the alternating response, from low to high values, appears to be identical (within the noise of the EP measurements) in the records for model and data. In addition, the average value of the

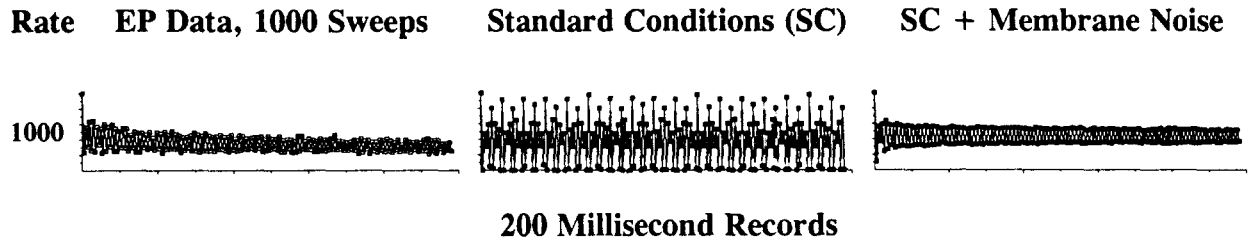

Fig. 25. EP magnitude data and model predictions for stimulation at 1000 pps. The EP magnitude data are from the combined records for the 1000 pps condition (1000 sweeps for each pulse polarity; see Fig. 13). Model predictions are the same as those in Fig. 24.

alternating response after the second pulse is approximately the same for the two records over the first 50 ms.

One detail in the data not predicted by the model is the magnitude of the response to pulse 2 for stimulation at high rates. The response to pulse 2 for the 800 and 1000 pps conditions is somewhat greater in the data than in the model (see Figs. 26 and 27). A possible explanation for this discrepancy is that the present model does not include a description of nonlinear node dynamics for neural membranes (only the passive component of membrane dynamics is included, see Fig. 2). In particular, incorporation of sodium channel dynamics may produce a better agreement between model and data. The nonlinear response of the membrane with sodium channels can act to integrate subthreshold stimuli. This can produce reductions in thresholds to closely spaced stimuli, as observed in psychophysical measures with implant patients (see, e.g., Shannon, 1983; Shannon, 1985; Eddington et al., 1994), but not predicted by the present model. Integration across pulses may allow neurons receiving relatively weak stimuli to reach a depolarization required for discharge, even though one pulse alone would not be sufficient to produce a discharge. Thus, some neurons might not respond to an initial pulse in a train of pulses, but could respond to a subsequent pulse, through temporal integration. Such "additional" responses could increase the overall number of neurons participating in the response to the second pulse for the 800 and 1000 pps conditions reviewed above. This would in turn produce an increment in the magnitude of the EP for pulse 2. We plan to incorporate a description of sodium channel dynamics in a new version of the population model to evaluate this possibility.

Results from the studies to evaluate effects of changes in each single parameter of the present model are presented in Appendix 2. In general, the effects produced by the addition of membrane noise, as described above, were greater (usually much greater) than the effects produced by changes in the values of any other single parameter.

SAM Pulse Trains

Further comparisons of model predictions and data for SAM pulse trains are presented in Table 2 and Figs. 28 through 31. Table 2 shows correlations between model and data for three variations of the

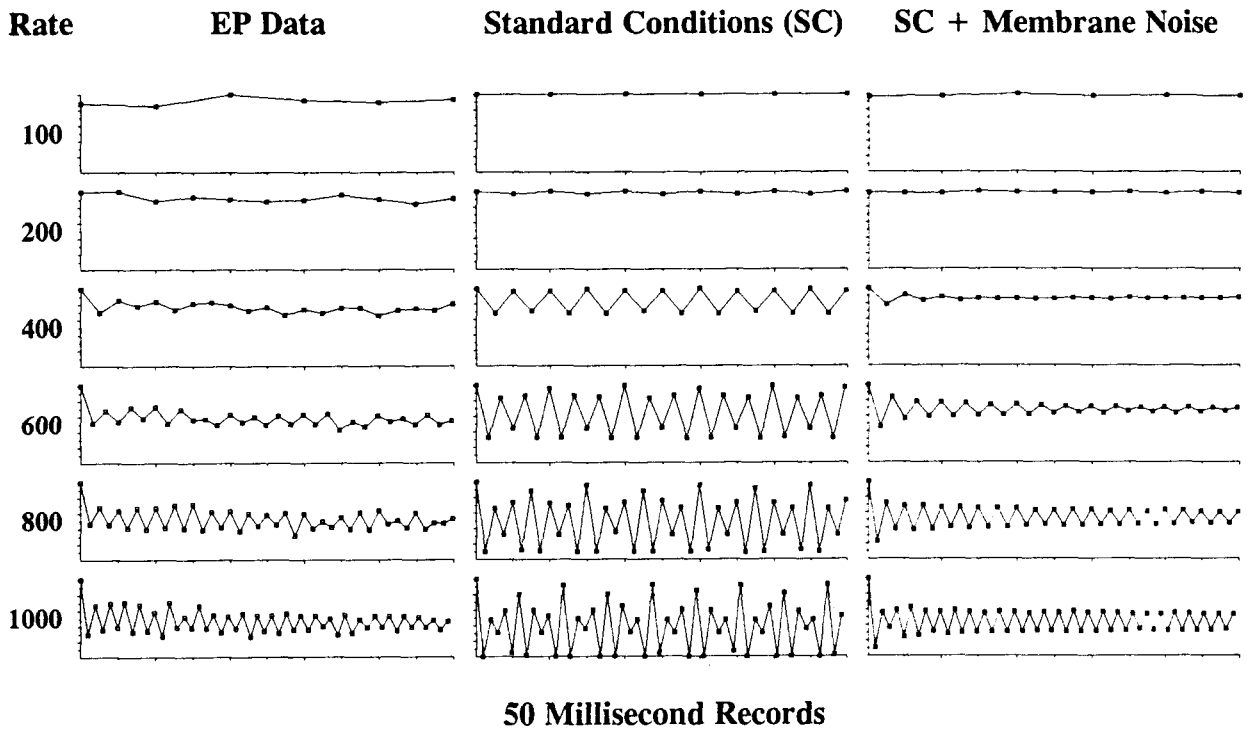

Fig. 26. Early portions of the records shown in Fig. 24, expanded time scale.

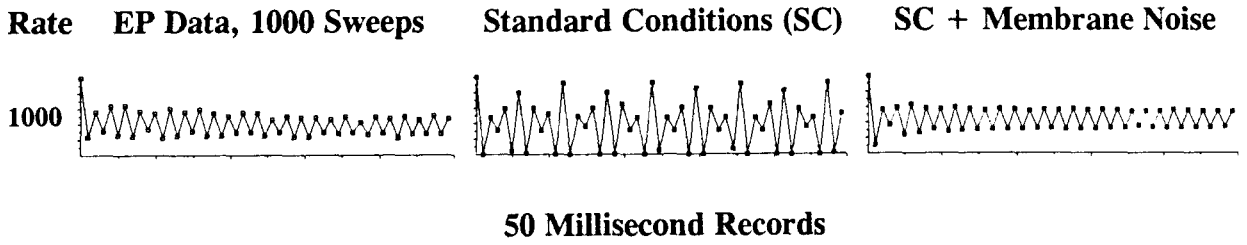

Fig. 27. Early portions of the records shown in Fig. 25, expanded time scale.

model. The data are the derived EP magnitudes from the recordings presented in Figs. 15 and 16. The model variations include the model with the standard conditions, the model with the new recovery function for each neuron, and a model with a low level of membrane noise (corresponding to level 1 in Fig. 8) added to the standard conditions. As indicated in the Table, agreement between model predictions and data is improved with the addition of membrane noise, but largely unchanged with substitution of the new recovery function. With the addition of membrane noise, the poorest correlation

Table 2. Pearson correlations between evoked potential magnitudes and model predictions for the indicated conditions of stimulation with 200 ms SAM pulse trains (corresponding to r_{200} values in Figs. 17 to 20).

		Model Conditions		
Modulation Freq (Hz)	Carrier Rate (pps)	Standard Conditions	New Recovery Function	Std Conditions Plus Noise
50	250	.997	.992	.999
100	250	.999	.999	.999
50	500	.975	.976	.986
100	500	.884	.864	.942
150	500	.976	.973	.977
200	500	.998	.997	.998
50	1000	.927	.912	.954
100	1000	.918	.934	.929
150	1000	.953	.961	.959
200	1000	.948	.949	.989
300	1000	.986	.982	.991
400	1000	.986	.967	.968
	Mean =	.962	.959	.974

between model and data improves from 0.88 to 0.93. Other levels of membrane noise may produce even better fits between model and data. Note that the best fit for pulse trains was obtained with a noise level of 2. Simulations are in progress to evaluate the possibility of a better fit between model and data for SAM pulse trains using various other levels of membrane noise (0.5, 1.5, 2.0, 2.5 and 3.0).

Comparisons of actual versus predicted EP magnitudes for the model with membrane noise (at the 1.0 level) are presented in Figs. 28-31. In general, this model predicts a better neural following of the stimuli than the model without noise (Figs. 17-20). This better following is more consistent with the EP data. For the carrier rate of 500 pps and the modulation frequency of 100 Hz, for example (Fig. 29), the model with membrane noise predicts a much larger response to the fourth pulse than the model without membrane noise (Fig. 18). However, the EP magnitude for that pulse still is greater than the magnitude predicted by the model with membrane noise. Further improvements in the agreement between model and data may be produced by the addition of higher levels of membrane noise in the

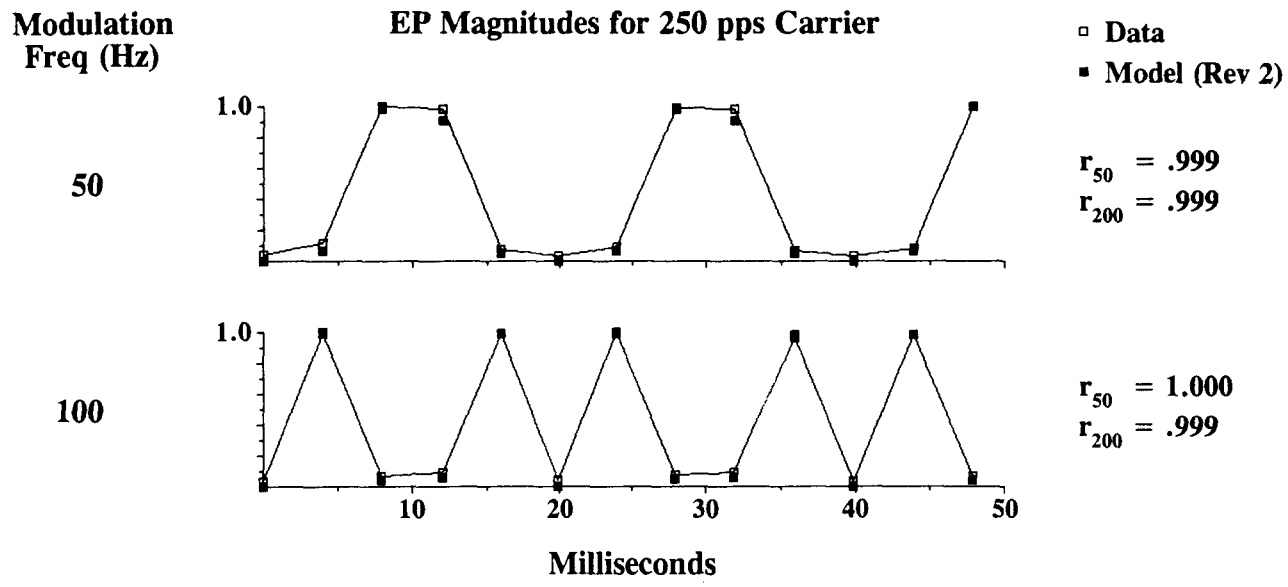

Fig. 28. Comparison of evoked potential magnitudes with model predictions of the number of neurons responding to each pulse in sinusoidally amplitude modulated (SAM) trains of pulses. The EP magnitudes were derived from the recordings presented in Fig. 15, for the carrier rate of 250 pps. The standard conditions plus membrane noise (at a level of 1.0) were used for the model simulations, with the same ratios of pulse amplitudes used in the experiment of Fig. 15. Data and model are compared for the first 50 ms of the stimuli in this illustration. Pearson correlations of data and model are presented to the right for the first 50 ms (r_{50}) and for the entire 200 ms (r_{200}) for each condition.

model or by the incorporation of sodium channel dynamics in the description of single neuron behavior.

Discussion

The results reviewed above have important implications for stimulus representation and coding with cochlear implants. The patterns of recorded responses to pulse trains show an inability of the auditory nerve to follow pulses with equal magnitudes of EPs at rates much above 400 pps. The rate at which the nerve begins to depart from an equal-magnitude following of a repetitive stimulus may correspond to the point at which judgments of pitch begin to asymptote with further increases in rate of stimulation. If such a connection exists between the peripheral representation and pitch judgments, then it may be possible to repair the representation so that higher pitch percepts can be produced in response to higher frequencies of stimulation. We will suggest several strategies for such repair in a future report. Improvements in temporal representations with implants could produce large improvements in speech recognition, by providing access to a greater range of frequencies (or modulation frequencies) for implant users.

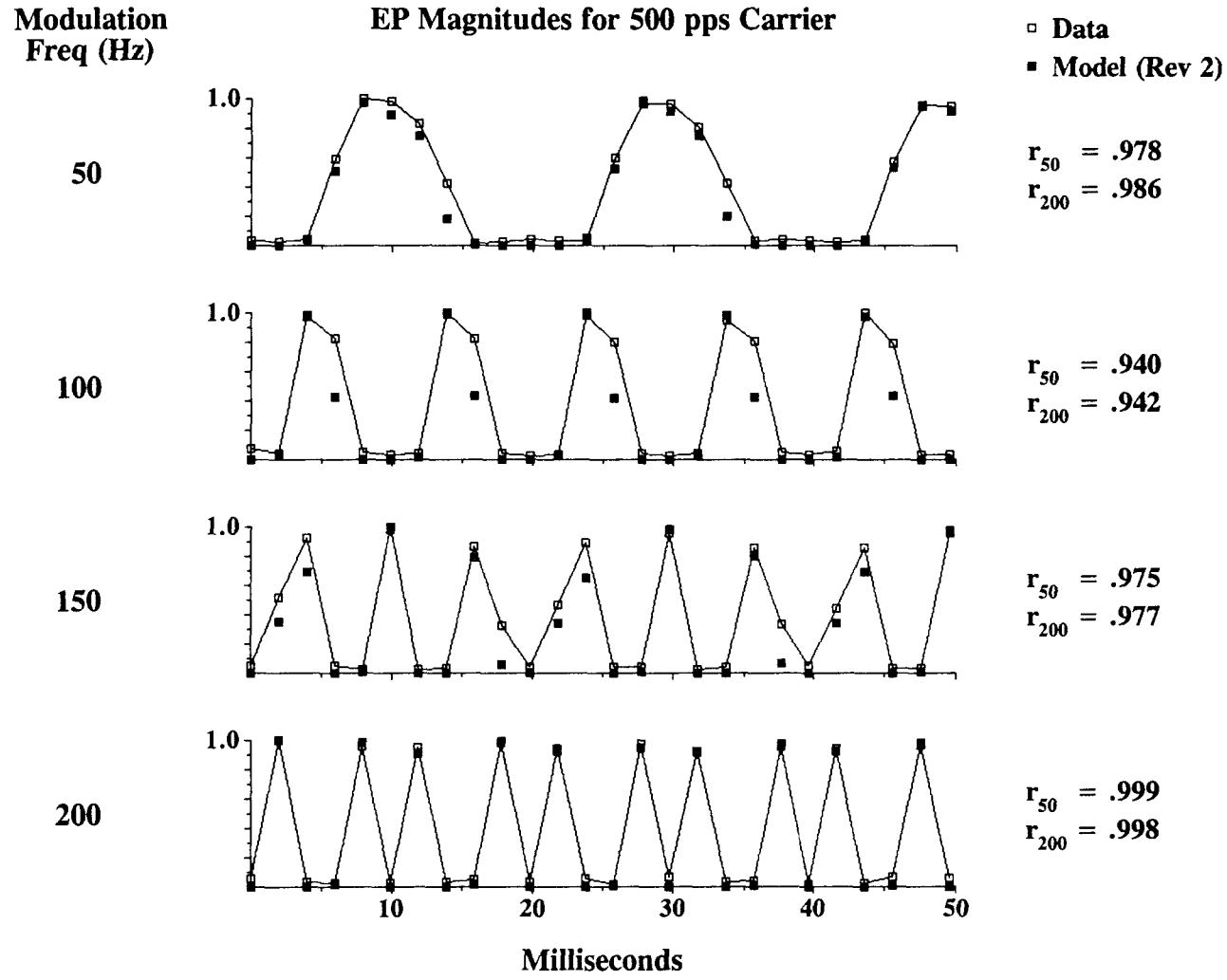

Fig. 29. Comparison of evoked potential magnitudes with model predictions of the number of neurons responding to each pulse in sinusoidally amplitude modulated (SAM) trains of pulses. The EP magnitudes were derived from the recordings presented in Fig. 15, for the carrier rate of 500 pps. Parameters for model simulations and organization of the panels are the same as those in Fig. 28.

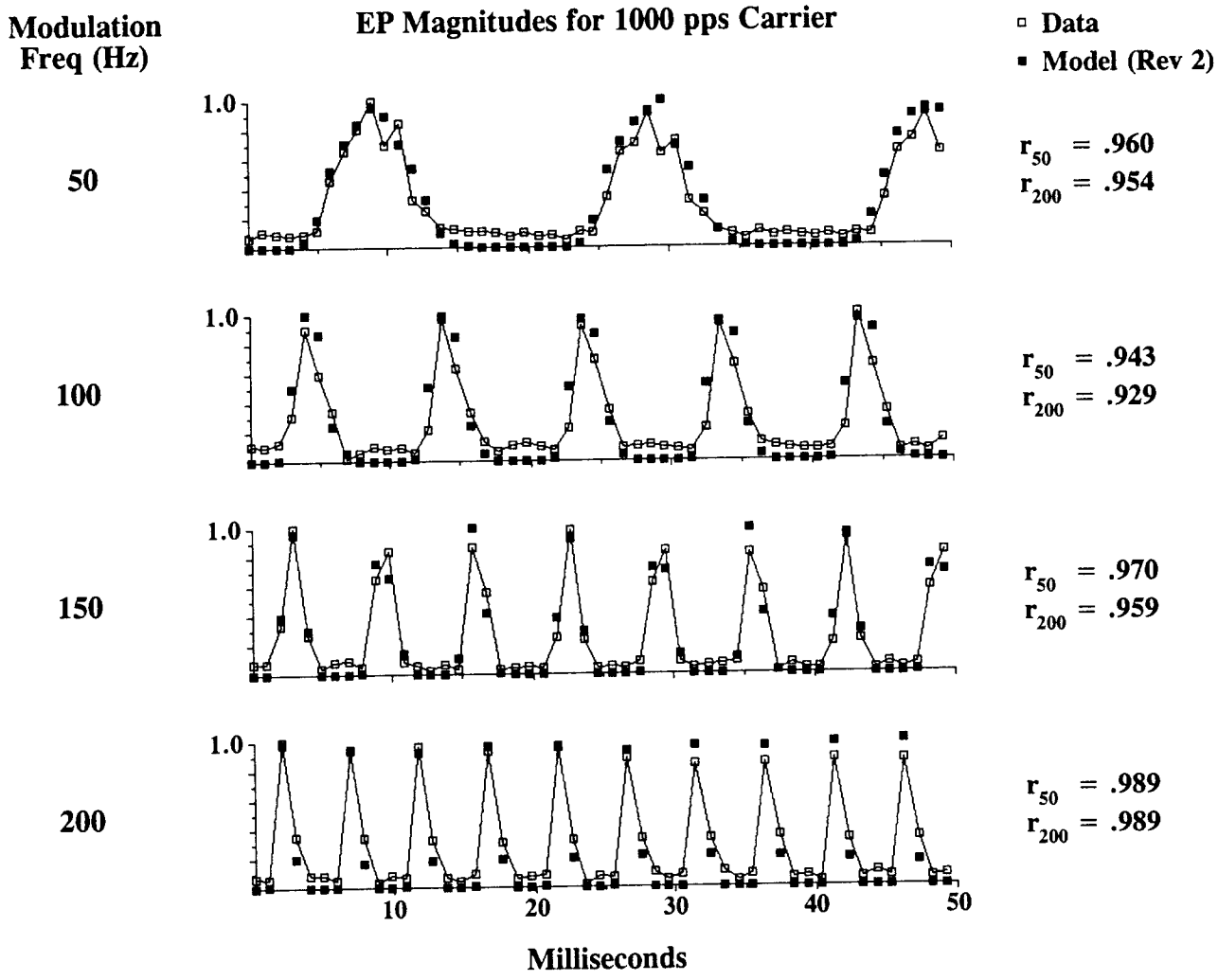

Fig. 30. Comparison of evoked potential magnitudes with model predictions of the number of neurons responding to each pulse in sinusoidally amplitude modulated (SAM) trains of pulses. The EP magnitudes were derived from the recordings presented in Fig. 15, for the carrier rate of 1000 pps. Parameters for model simulations and organization of the panels are the same as those in Fig. 28.

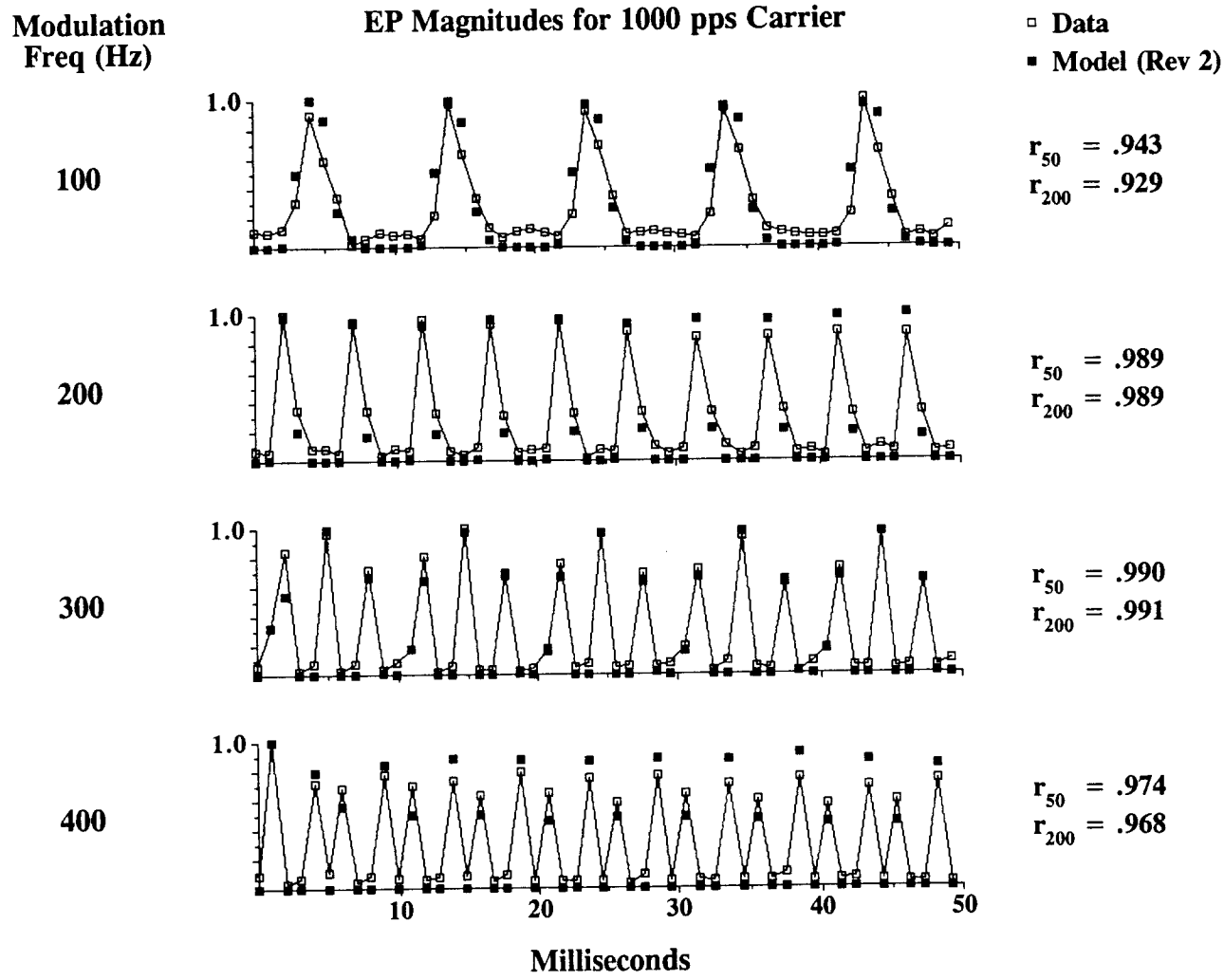

Fig. 31. Comparison of evoked potential magnitudes with model predictions of the number of neurons responding to each pulse in sinusoidally amplitude modulated (SAM) trains of pulses. The EP magnitudes were derived from the recordings presented in Fig. 16, for the carrier rate of 1000 pps. Parameters for model simulations and organization of the panels are the same as those in Fig. 28.

As noted in Fig. 8 and the accompanying discussion, membrane or synaptic noise may have profound effects on representations of repetitive stimuli. This not only has important implications for modeling and understanding population responses with implants, but also raises a caution in interpreting results from prior studies of electrical stimulation with intact animal preparations. In particular, even if the possibility of electrophonic responses can be eliminated in such studies, the intact synapses between cochlear hair cells and auditory neurons still will release chemical transmitter substance into the cleft. This release in the absence of stimulation produces a high level of noise at the synaptic endings of auditory neurons, high enough to produce spontaneous activity (corresponding to level 3 in Fig. 8). The presence of such noise can alter greatly the predicted patterns of response to repetitive stimuli. Thus, results reported for intact animals may differ markedly from results obtained with deafened animals, particularly for stimuli presented at relatively high rates or frequencies.

By the same token, patients with some residual function of the synaptic machinery may exhibit better abilities to scale and discriminate frequencies than patients without such function (see also, Risberg et al., 1990). Similarly, patients with relatively high levels of neural membrane noise may have better frequency scaling and discrimination abilities than patients with low levels of membrane noise. It may be possible to develop ways to measure indirectly levels of membrane or synaptic noise in implanted patients, and it may also be possible to develop ways to exploit whatever level of noise remains. Ironically, presence of noise in the system may have a salutary effect on stimulus representations with implants.

Modeling studies to date indicate the likely importance of some level of neural membrane noise in replicating the patterns of intracochlear EPs recorded from human subjects. With the addition of such noise, good agreement between model and data has been achieved for results from one subject (SR2). Studies are in progress to evaluate the predictive power of the model for results from the other subjects. We anticipate that different adjustments of model parameters will be required to achieve the best fits between model and data for different subjects. If so, the parameters for each subject may provide an indication of the status of that person's auditory nerve, perhaps even as a function of position within the cochlea (from comparisons between model and data for stimulation of different electrodes).

Further improvements in the accuracy of model predictions may be obtained by (a) inclusion of sodium channel dynamics in the neural membrane description and (b) inclusion of a model element to account for long-term accommodation in neural responses. We plan to evaluate these possibilities in future work.

Results from the studies with SAM pulse trains show that the carrier rate must be substantially higher than the highest modulation frequency, for a smooth and unambiguous representation of the modulation waveform. For modulation frequencies near the Nyquist criterion (½ the carrier rate) the predicted and recorded patterns of neural response indicate only crude and ambiguous representations of the modulation waveform. When the modulation frequency exceeds the Nyquist criterion, aliasing occurs, producing misrepresentations of the modulation waveform. These findings have implications for the design of speech processors that use modulated pulse trains as stimuli. For example, effects of aliasing can be anticipated with the 250 pps carrier and maximum modulation frequency of 200 Hz used in the

"Spectral Maxima Sound Processor" and the related "Spectral Peak" (SPEAK) processor developed in Australia (see McDermott et al., 1992, and McKay et al., 1991, for descriptions of the SMSP). In general, our results suggest that the carrier rate should be at least four to five times the highest frequency in the modulation waveforms for a good representation of those waveforms with implants. We note that Busby et al. [1993] have offered this same suggestion, based on results from their psychophysical studies with patients using the clinical Nucleus device.

V. Plans for the Next Quarter

Our plans for the next quarter include the following:

1. Initial studies with Ineraid subject SR14 and continued studies with subject SR2. The studies with SR14 will include measures of intracochlear evoked potentials for a variety of stimuli, measures of speech reception with CIS processors using different pulse durations and rates, and measures of speech reception with single-channel processors. Studies with SR2 will include additional measures of intracochlear EPs and further investigation of complex tone perception with cochlear implants.
2. Initial studies with the first patient in the Nucleus percutaneous series, NP-1. The studies will include evaluations of CIS and *spectral peak* (SPEAK) processing strategies.
3. Continued interaction with the groups at the Massachusetts Eye & Ear Infirmary in Boston and the Hôpital Cantonal Universitaire in Geneva, Switzerland, to develop a portable processor for use in research studies.
4. A site visit by Terry Hambrecht and Bill Heetderks.
5. Presentation of project results in invited lectures at the *127th Meeting of the Acoustical Society of America* (Cambridge, MA, June 8).
6. Continued preparation of manuscripts for publication.

VI. Acknowledgments

We thank the subjects of the studies described in this report for their enthusiastic participation and generous contributions of time. The initial design of the model of neural population responses to intracochlear electrical stimulation was supported by NIH project N01-NS-3-2356, "Speech Processors for Auditory Prostheses," and the subsequent development and application of the model, as described in this report, was supported under Project IV of NIH Program Project Grant P01-DC00036, "Mechanisms of Intracochlear Electrical Stimulation." The psychophysical and evoked potential studies presented in this report were supported by the present project, N01-DC-2-2401, "Speech Processors for Auditory Prostheses."

VII. References

Brown CJ, Abbas PJ (1990a) Electrically evoked whole-nerve action potentials: Data from human cochlear implant users. J. Acoust. Soc. Am. 88: 1385-1391.

Brown CJ, Abbas PJ (1990b) Electrically evoked whole-nerve action potentials: Parametric data from the cat. J. Acoust. Soc. Am. 88: 2205-2210.

Busby PA, Tong YC, Clark GM (1993) The perception of temporal modulations by cochlear implant patients. J. Acoust. Soc. Am. 94: 124-131.

Clopton BM, Glass I (1984) Unit responses at cochlear nucleus to electrical stimulation through a cochlear prosthesis. Hearing Res. 14: 1-11.

Dynes SBC, Delgutte B (1992) Phase-locking of auditory-nerve discharges to sinusoidal electric stimulation of the cochlea. Hearing Res. 58: 79-90.

Eddington DK, Rubinstein JT, Dynes SBC (1994) Forward masking during intracochlear electrical stimulation: Models, physiology, and psychophysics. J. Acoust. Soc. Am. 95: 2904.

Glass I (1985) Responses of cochlear nucleus units to electrical stimulation through a cochlear prosthesis: Channel interaction. Hearing Res. 17: 115-126.

Hartmann R, Klinke R (1990) Response characteristics of nerve fibers to patterned electrical stimulation. In: Miller JM, Spelman FA (Eds), *Cochlear Implants: Models of the Electrically Stimulated Ear*. New York: Springer-Verlag, pp. 135-160.

Hartmann R, Topp G, Klinke R (1984) Discharge patterns of cat primary auditory fibers with electrical stimulation of the cochlea. Hearing Res. 13: 47-62.

Hill AV (1936) Excitation and accommodation in nerve. Proc. Royal Soc. 119: 305-355.

Javel E (1990) Acoustic and electrical encoding of temporal information. In: Miller JM, Spelman FA (Eds), *Cochlear Implants: Models of the Electrically Stimulated Ear*. New York: Springer-Verlag, pp. 247-296.

Javel E, Tong YC, Shepherd RK, Clark GM (1987) Responses of cat auditory nerve fibers to biphasic electrical current pulses. Ann. Otol. Rhinol. Laryngol. 96, Suppl. 128: 26-30.

Kiang NYS, Moxon EC (1972) Physiological considerations in artificial stimulation of the inner ear. Ann. Otol. 81: 714-730.

Knauth M, Hartmann R, Klinke R (1994) Discharge pattern in the auditory nerve evoked by vowel stimuli: A comparison between acoustical and electrical stimulation. Hearing Res. 74: 247-258.

McDermott HJ, McKay CM, Vandali AE (1992) A new portable sound processor for the University of Melbourne/Nucleus Limited multielectrode cochlear implant. J. Acoust. Soc. Am. 91: 3367-3391.

McKay C, McDermott H, Vandali A, Clark G (1991) Preliminary results with a six spectral maxima sound processor for the University of Melbourne/Nucleus multiple-electrode cochlear implant. J. Otolaryngol. Soc. Australia 6: 354-359.

McNeal DR (1976) Analysis of a model for excitation of myelinated nerve. IEEE Trans. on Biomedical Engineering. 23: 329-337.

Parkins CW (1989) Temporal response patterns of auditory nerve fibers to electrical stimulation in deafened squirrel monkeys. Hearing Res. 41: 137-168.

Parkins CW, Colombo J (1987) Auditory-nerve single-neuron thresholds to electrical stimulation from scala tympani electrodes. Hearing Res. 31: 267-286.

Risberg A, Agelfors E, Lindström B, Bredberg G (1990) Electrophonic hearing and cochlear implants. Acta Oto-Laryngol., Suppl. 469: 156-163.

Shannon RV (1983) Multichannel electrical stimulation of the auditory nerve in man. I. Basic Psychophysics. Hearing Res. 11: 157-189.

Shannon RV (1985) Threshold and loudness functions for pulsatile stimulation of cochlear implants. Hearing Res. 18: 135-143.

van den Honert C, Stypulkowski PH (1984) Physiological properties of the electrically stimulated auditory nerve. II. Single fiber recordings. Hearing Res. 14: 225-243.

van den Honert C, Stypulkowski PH (1987a) Single fiber mapping of spatial excitation patterns in the electrically stimulated auditory nerve. Hearing Res. 29: 195-206.

van den Honert C, Stypulkowski PH (1987b) Temporal response patterns of single auditory nerve fibers elicited by periodic electrical stimuli. Hearing Res. 29: 207-222.

Verveen AA (1962) Fibre diameter and fluctuation in excitability. Acta Morphologica Neerlando-Scandinavica. 5: 79-85.

Verveen AA, Derksen HE (1968) Fluctuation phenomena in nerve membrane. Proc. IEEE 56: 906-916.

Wilson BS (1986) Ensemble models of neural discharge patterns evoked by intracochlear electrical stimulation. Invited speaker presentation at the *IUPS Satellite Symposium on Advances in Auditory Neuroscience*, San Francisco, CA, July 8-11, 1986.

Wilson BS, Finley CC (1986) Latency fields in electrically-evoked hearing. ARO Abstracts, 9th Midwinter Res. Conf., St. Petersburg, FL, pp. 170-171.

Wilson BS, Finley CC, Lawson DT (1985) Speech processors for auditory prostheses. *Eighth Quarterly Progress Report*, NIH project N01-NS-3-2356. Bethesda, MD: Neural Prosthesis Program, National Institutes of Health.

Wilson BS, Finley CC, Lawson DT, Wolford RD, Eddington DK, Rabinowitz WM (1991) Better speech recognition with cochlear implants. Nature 352: 236-238.

Wilson BS, Lawson DT, Zerbi M, Finley CC (1994) Recent developments with the CIS strategies. In IJ Hochmair-Desoyer and ES Hochmair (Eds.), *Advances in Cochlear Implants*, International Interscience Seminars Series, Manz, Vienna, pp. 103-112.

Appendix 1

Design for an inexpensive but effective cochlear implant system:
Recommendations of an expert panel from the *1993 Zhengzhou International Symposium on Electrical Cochlear Hearing and Linguistics*

[paper accepted for presentation at the *International Cochlear Implant, Speech and Hearing Symposium*, Melbourne, Australia, October 23-28, 1994]

DESIGN FOR AN INEXPENSIVE BUT EFFECTIVE COCHLEAR IMPLANT SYSTEM: RECOMMENDATIONS OF AN EXPERT PANEL FROM THE *1993 ZHENGZHOU INTERNATIONAL SYMPOSIUM ON ELECTRICAL COCHLEAR HEARING AND LINGUISTICS*

GE Loeb[1], RV Shannon[2], BS Wilson[3], F-G Zeng[2] and S Rebscher[4]*

[1]Queen's University, Kingston, Ontario, CANADA
[2]House Ear Institute, Los Angeles, CA, USA
[3]Research Triangle Institute, Research Triangle Park, NC, USA
[4]University of California at San Francisco, San Francisco, CA, USA

*Authors listed in order of a random draw

A design for a low cost, but effective, cochlear implant system is described. Such a system may be especially appropriate for widespread use in developing countries. The system includes a speech processor, four pairs of transmitting and receiving coils, and an electrode array with four monopolar electrodes. All implanted components are passive, reducing to a minimum the complexity of manufacture and allowing high reliability. A transcutaneous link is used to minimize the possibility of infection, which in turn should minimize medical costs in maintaining device function. The electrode array has a mechanical memory to position the electrodes in close proximity to the inner modiolar wall of the scala tympani. A reference electrode is implanted in the temporalis muscle. A four-channel *continuous interleaved sampling* (CIS) strategy is used for the speech processor. The speech processor and transcutaneous link have been evaluated in preliminary tests with a patient implanted with the Ineraid electrode array and percutaneous connector. A prototype of the link, consisting of four pairs of transmitting and external receiving coils, was used, with the outputs of the receiving coils routed to the apical four electrodes of the Ineraid array via the percutaneous connector. The tests included identification of 24 consonants in an /a/-consonant-/a/ context. The subject scored $89 \pm 2\%$ correct with a standard laboratory implementation of the four-channel CIS processor using current-controlled stimuli and $92 \pm 1\%$ correct with the prototype system. These scores show that use of the coils does not degrade performance. Results from other studies indicate that many patients can achieve high levels of speech recognition with CIS processors and monopolar electrodes. The system recommended here includes those components along with (a) an electrode array that may improve placement of the intracochlear contacts and (b) a simple transcutaneous link that does not degrade performance.