3

Ninth Quarterly Progress Report

August 1 through October 31, 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 subject SR2, primarily recording intracochlear evoked potentials for a wider range of stimuli (August 15 through September 2).
2. Initial studies with the second patient in the Nucleus percutaneous series, NP-2. Studies included evaluations of CIS and clinical *spectral peak* (SPEAK) processing strategies.
3. Continued analysis of speech reception and evoked potential data from prior studies.
4. Presentation of project results at the *25th Annual Neural Prosthesis Workshop* (Bethesda, MD, October 18-21) and at the *International Cochlear Implant, Speech and Hearing Symposium* (Melbourne, Australia, October 23-28).
5. Continued preparation of manuscripts for publication.

In the Seventh Quarterly Progress Report for this project we described limitations and distortions in temporal representations with cochlear implants, as demonstrated by recordings of intracochlear evoked potentials. In the present report we describe strategies to reduce or repair such deficits. Results from the studies indicated in points 1 and 2 above will be presented in future reports.

II. Strategies for the Repair of Deficits in Temporal Representations with Implants

Results from the evoked potential studies described in the Seventh Quarterly Progress Report for this project have demonstrated limitations in the ability of the auditory nerve to follow electrical pulses at rates much above 200/s. In addition, the recorded responses to sinusoidally amplitude modulated (SAM) pulse trains have demonstrated various distortions in the neural representation of the modulation waveform, depending on carrier rate and modulation frequency.

In the present report we describe strategies to reduce or repair distortions in the neural encoding of repetitive electrical stimuli. Improvements in the fidelity of neural encoding might allow patients to perceive a wider range of frequencies via their implants. Such improvements also might reduce uncertainties in judgments about certain stimuli, e.g., the perceptual ambiguities noted in QPR 7 for SAM pulse trains (see pp. 23-26 of QPR 7).

We note that temporal representations with implants are quite crude and limited in comparison with the representations found in normal hearing. In the normal case, a much higher level of asynchrony and stochastic independence is found among neurons in the auditory nerve (see, e.g., Abbas, 1993). This allows the nerve to follow acoustic stimulation frequencies up to several kHz in the population response to repetitive stimuli (e.g., Young and Sachs, 1979; also see Figure 8 in QPR 7).

It may well be that improvements in temporal representations with implants could produce large improvements in speech reception. In addition, improvements in temporal representations might produce improvements in the "naturalness" of percepts for implant users, through a closer approximation to patterns of response found in the normal auditory nerve.

Possible ways to improve frequency following in electrically stimulated auditory nerves have been developed through use of our population model, described in QPR 7. These have included (a) adjusting pulse amplitudes so that an equal number of spikes is produced by each pulse in a train of pulses; (b) sharing stimulus pulses across electrodes, to distribute demand in conveying temporal information across subpopulations of neurons; (c) using stimulus waveforms that exploit the presence of membrane or synaptic noise; and (d) using high carrier rates for modulated pulse trains.

Adjustment of Pulse Amplitudes

Improvement through the adjustment of pulse amplitudes is illustrated in Figure 1. Patterns of temporal response predicted by the population model are shown for our standard conditions without membrane noise (see QPR 7). The response for a train of constant amplitude pulses is presented in the left column and the response for a train of pulses with adjusted pulse amplitudes is presented in the right column. The pulse rate is 1000/s for both columns.

Stimulation with a train of identical pulses produces a large number of spikes in response to the first pulse, none for the second pulse, and a variable number for subsequent pulses, as previously described in QPR 7.

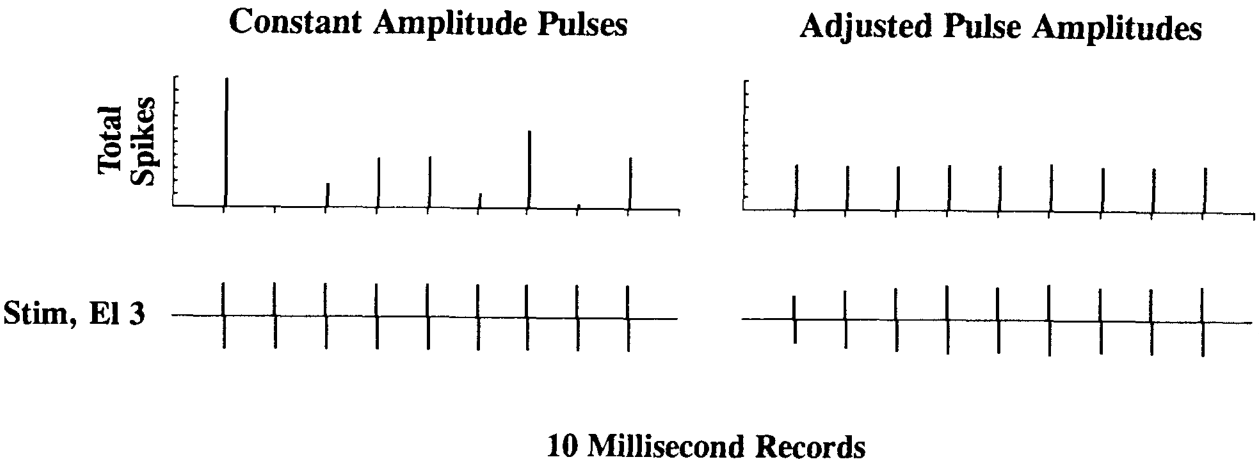

Figure 1. Predicted patterns of neural response to a train of constant amplitude pulses (left column) and to a train of pulses whose amplitudes have been adjusted to produce an equal number of spikes for each pulse (right column). The stimulus waveform is presented in the bottom panel of each column and the predicted patterns of neural response in the top panel.

In contrast, a uniform pattern of responses across sequential pulses can be obtained through (relatively small) adjustments of pulse amplitudes. As indicated in the right column of Figure 1, the amplitude of pulse 1 is reduced so that the average of the responses for all pulses in the left column is produced. This leaves a greater number of neurons available for response to the second pulse. The amplitude of that pulse is adjusted to produce the same number of spikes, as are the amplitudes of all subsequent pulses. The result is a set of afferent volleys from the auditory nerve at the rate of pulses in the stimulus and with equal numbers of spikes in each volley. If the central auditory system can make use of such an input, then pitch percepts produced by stimuli with adjusted pulse amplitudes might be higher than pitch percepts for stimuli without the adjustments. Also, the percepts with the adjusted pulse amplitudes may be purer inasmuch as the spectral structure of the afferent input to the central system is much simpler for the adjusted amplitudes case.

As a first step in evaluating this strategy, we have demonstrated that patterns of intracochlear EPs can be altered through manipulations in pulse amplitudes. Results for Ineraid subject SR2 are presented in Figures 2 and 3. In Figure 2 the amplitude of the first pulse was set at 1.0, 0.5, 0.7 or 0.8 times the amplitude of the (identical) remaining pulses. The pulse rate was 1000/s. The control condition, with all pulses of equal amplitude, produces the alternating pattern of EPs described before for this subject (e.g., see Figure 9 in QPR 7). Reductions in the amplitude of the first pulse produce reductions in the magnitude of the EP following that pulse. No EP is detected for the 0.5 amplitude pulse (for the number of sweeps used in the present recordings), and a relatively small EP is seen for the 0.7 amplitude pulse. The 0.8 amplitude pulse produces an EP similar in magnitude to the grand average of

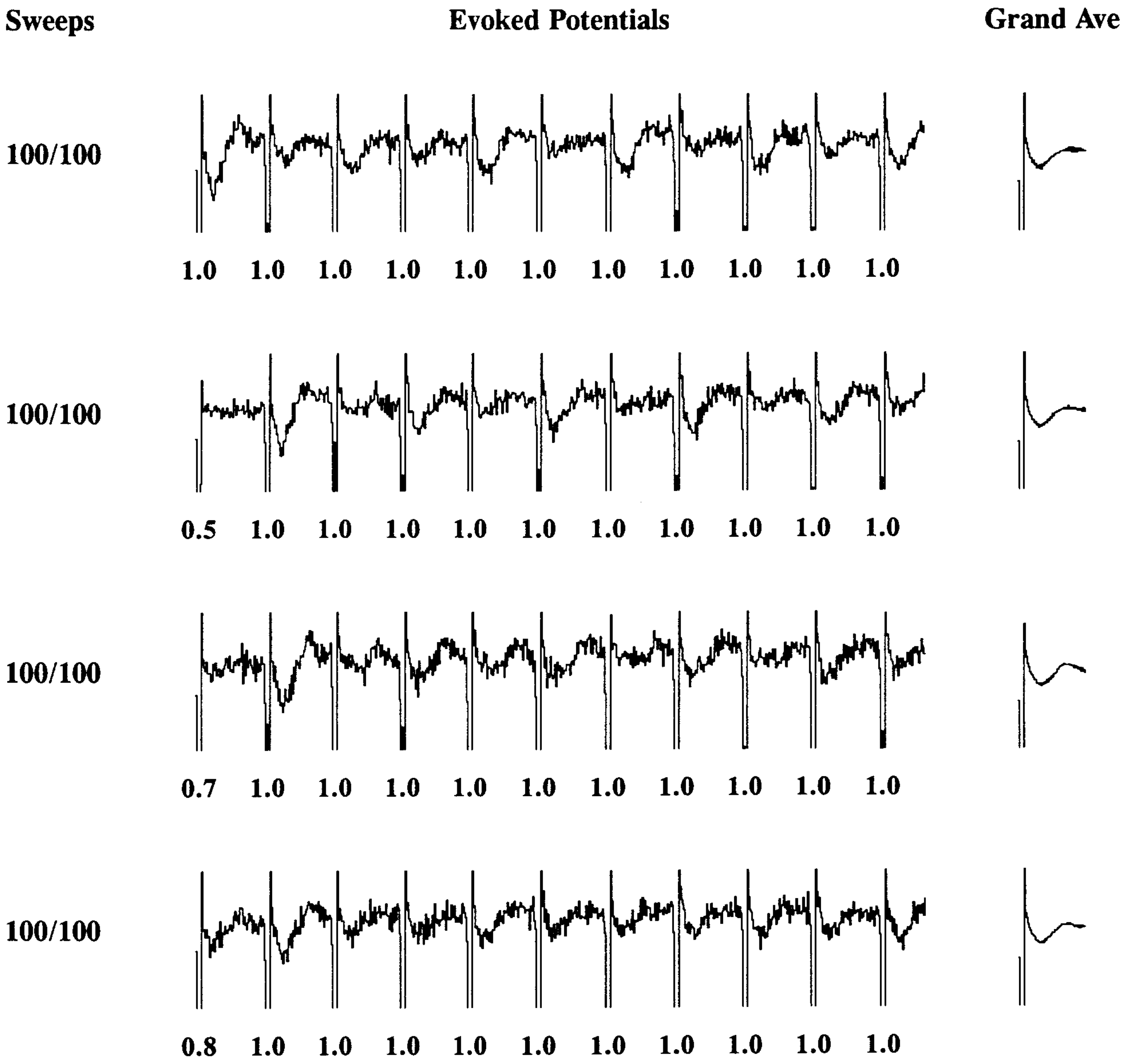

Figure 2. Recordings of intracochlear evoked potentials for Ineraid subject SR2. Stimuli included 16.4 μs/phase pulses presented at 1000 pps to electrode 3. The number beneath each pulse artifact indicates the relative amplitude of the corresponding pulse. Maximum pulse amplitude was 750 μA. The number of sweeps used in the recordings for each polarity of pulses is indicated in the left column for each condition. The average of EPs following each pulse for a given condition is shown in the right column, under the heading of "Grand Ave." Potentials were recorded differentially between electrode 4 and the ipsilateral mastoid.

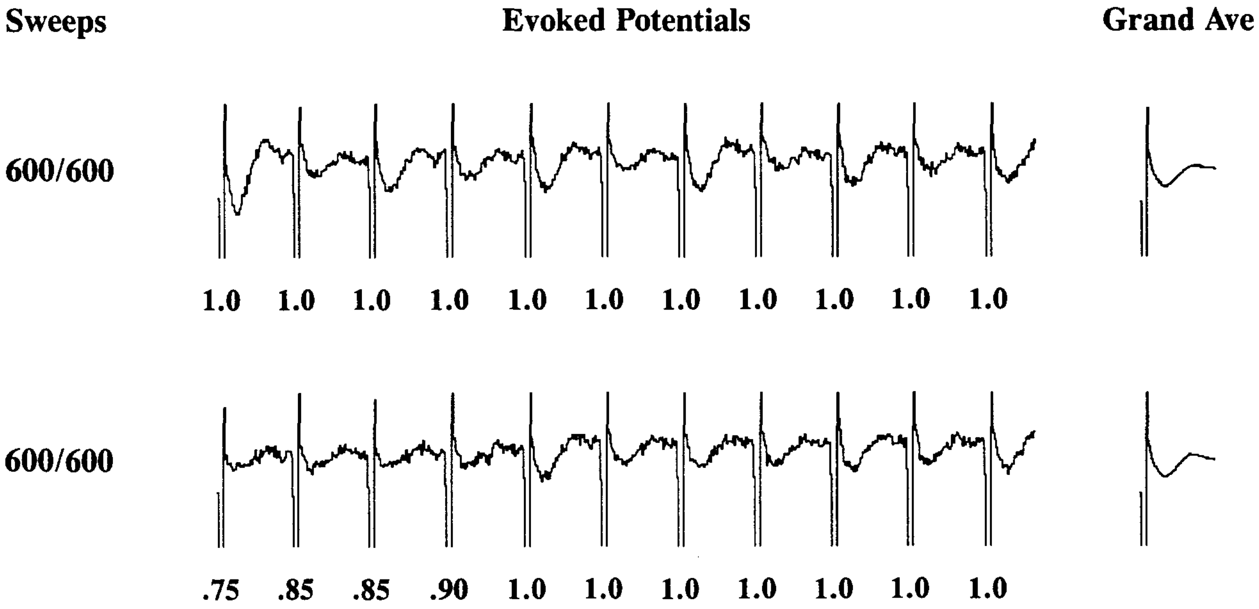

Figure 3. Recordings of intracochlear evoked potentials for Ineraid subject SR2. Organization for the figure is the same as that for Figure 2. The present figure shows results obtained with adjustment of the first four pulses in the burst.

EPs, across the 200 pulses in each 200 ms burst of pulses. With reductions in the magnitude of the first EP, we would expect that fewer neurons would be in refraction at the time of the second pulse. This idea is supported by the relatively large magnitudes of EPs following the second pulse for the three conditions of reduced amplitudes for pulse 1.

In a second experiment, iterative adjustments were made in the amplitudes of the first four pulses in the burst to produce approximately equal magnitudes of EPs. The result is shown in Figure 3.

Ultimately, we will want to produce equal magnitudes of EPs across all pulses in the 200 ms burst. Once that is accomplished, we will be in a position to conduct psychophysical experiments to compare pitch percepts elicited with a train of constant amplitude pulses versus a train of pulses with adjusted amplitudes. We plan to conduct such experiments under support of a separate project. If the outcome is positive, i.e., if pitch percepts change in the anticipated ways, then we will begin work to implement the approach in a speech processor design, either in that separate project or the present project, depending on availability of investigator time and other resources.

We note that such a positive outcome from the psychophysical experiments would have important implications for basic auditory theory.

Sharing of Pulses Across Electrodes

The strategy of sharing pulses across electrodes is illustrated in Figure 4. The left column shows a control condition, as in Figure 1, with stimulation of only one electrode using a train of constant amplitude pulses. The right column shows the predicted pattern of neural response for pulses distributed across eight electrodes, spaced 4 mm apart. Electrode 1 receives the first pulse, electrode 4 the second, electrode 7 the third, and so on. The rate of stimulation on any one electrode is 125 pps. As indicated in the model prediction, the local neural population can follow this low rate with high uniformity from pulse to pulse. The aggregate rate across electrodes is 1000 pps. The afferent volley, summed across the entire auditory nerve, includes equal numbers of spikes at 1 ms intervals. If the central auditory system were to integrate this input spatially across the auditory nerve, then a percept corresponding to a 1000 pps stimulus might be produced.

To investigate this possibility, we conducted magnitude estimation experiments with two Ineraid patients (subjects SR2 and SR13) in which percepts elicited by stimulation of a single electrode were compared with percepts elicited by stimulation of multiple electrodes, in the "shared electrodes" paradigm outlined above. The subjects were asked to nominate a pitch between 1 and 100 for each stimulus. The stimuli included 200 ms bursts of pulses presented singly to electrode 3 at the rates of 66.7, 133.3, 200, 266.7, 400, 600 and 800 pps, or across three adjacent electrodes (electrodes 2, 3 and 4) at the aggregate rates of 200, 400, 600, 800, 1200, 1800 and 2400 pps. The results for subject SR2 are shown in Figure 5. The open squares show the mean of pitch judgments for stimulation of electrode 3 alone, and the triangles show the mean of pitch judgments for stimulation across electrodes 2, 3 and 4. The open triangles show those results for the aggregate rate across electrodes and the closed triangles show the results for the rate on any one of the three electrodes. The results for subject SR13 were essentially identical to those shown in Figure 5 for SR2.

As is evident from the figure, use of shared electrodes did not produce pitches higher than those elicited with stimulation of a single electrode. That is, the judgments for shared electrodes stimulation overlie the judgments for single electrode stimulation when the shared electrodes judgments are plotted according to the rate on any one electrode. The subjects apparently were able to separate the inputs from the different cochlear regions stimulated by the different electrodes in the shared electrodes paradigm in making judgments related to rate of stimulation. The central auditory system did not, at least in this case, integrate the different inputs spatially.

The 4 mm spacing of electrodes in the Ineraid implant may be too great to demonstrate such integration. Spatial integration of inputs by the central auditory system in normal hearing occurs within critical bands -- corresponding to approximately 1 mm regions along the cochlear partition -- for a wide range of tasks and behaviors.

In contrast to the present findings, a shared electrodes paradigm might produce higher pitch percepts if multiple electrodes could be placed within a critical band distance, and if those electrodes could stimulate different populations of neurons. As testing time permits, we hope to conduct studies with patients in our "Nucleus Percutaneous" series to evaluate what may be a closer approach to that situation. The electrodes in the Nucleus array are spaced 0.75 mm apart, just within a critical band

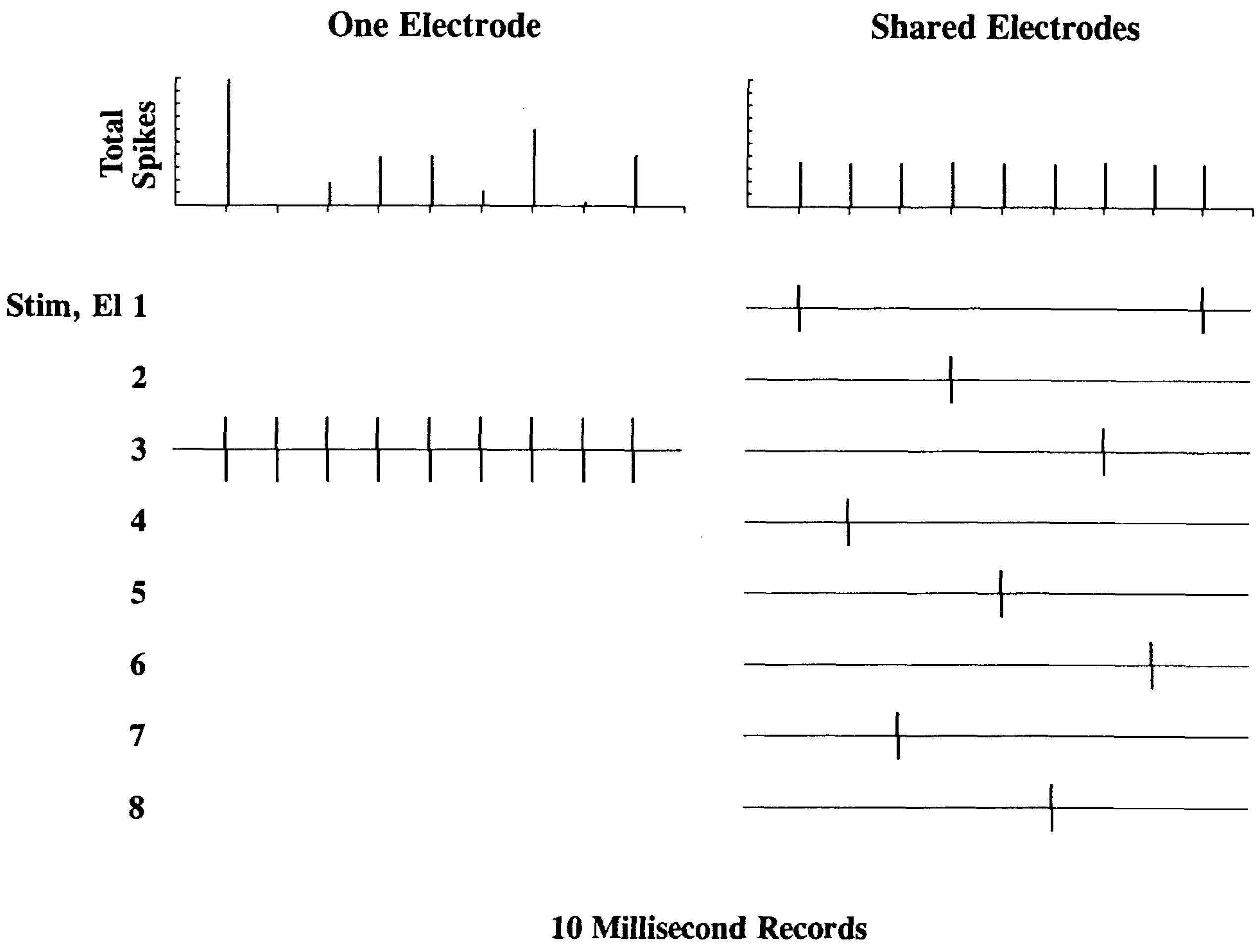

Figure 4. Predicted patterns of neural response to a train of constant amplitude pulses delivered to a single electrode (left column) and to interleaved trains of constant amplitude pulses delivered to eight electrodes (right column). The stimuli are presented in the bottom panels and predicted patterns of neural response in the top panels.

distance. If we identify a patient who can reliably discriminate adjacent electrodes on the basis of pitch judgments, then we will use those electrodes in a repetition of the experiment described above. The question will be whether the Nucleus subject perceives higher pitches with shared stimulation of the two adjacent electrodes than with stimulation of either electrode alone. A positive outcome would provide a strong incentive to develop new types of electrodes for cochlear implants, with a much greater number and density of contacts compared with present electrodes. Such an electrode might allow independent or quasi-independent stimulation of three or more populations of neurons within a critical band,

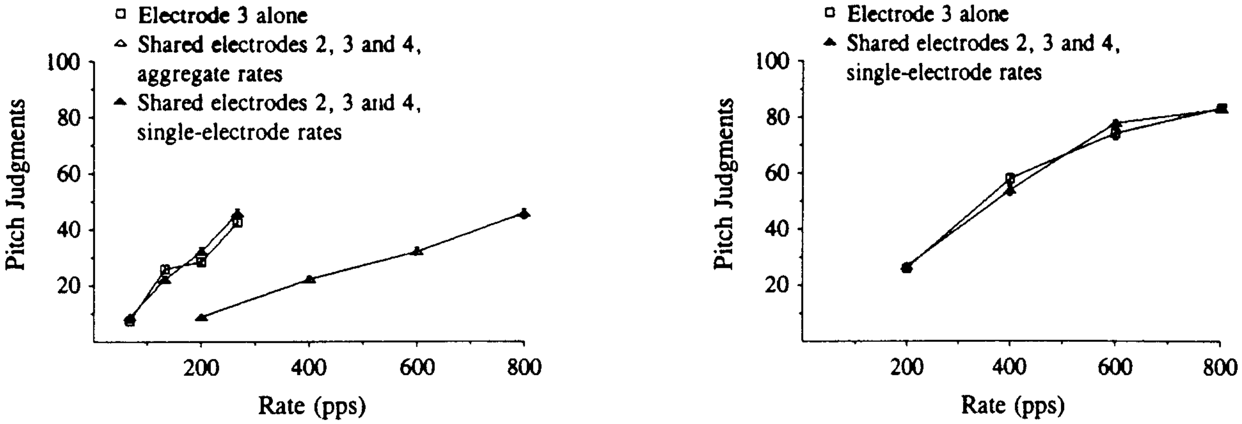

Figure 5. Results from a magnitude estimation experiment, in which the subject (SR2) was asked to nominate a pitch for each stimulus. The subject was instructed to nominate a low number for a low pitch and a high number for a high pitch. Stimuli included 200 ms bursts of 33 μs/phase pulses presented singly to electrode 3 at the rates of 66.7, 133.3, 200, 266.7, 400, 600 and 800 pps, or across three adjacent electrodes (electrodes 2, 3 and 4) at the aggregate rates of 200, 400, 600, 800, 1200, 1800 and 2400 pps. The pulse amplitude for each condition was adjusted to produce a most comfortable loudness (MCL) percept. Once the amplitudes for MCL percepts had been identified, the stimuli for the various conditions were played in sequence. Fine adjustments in pulse amplitudes were then made, and the sequence repeated, until the stimuli for all conditions were judged to be equally loud. In the experiments, the stimuli for the different conditions were presented in a randomized order according to the method of constant stimuli. The number of trials for each condition was 30. Results for conditions involving low rates are presented in the left panel and results for conditions involving high rates in the right panel.

especially if the contacts are placed in close proximity to the target spiral ganglion cells, to reduce overlaps in neural excitation fields associated with adjacent electrode positions.

Pending the availability of patient testing time and other resources, a positive outcome also would be pursued by us in subsequent studies with Nucleus percutaneous subjects. In particular, we would incorporate the shared electrodes paradigm into a speech processor design and determine whether shared stimulation of pairs of adjacent electrodes may provide an advantage in recognizing speech.

Selection of Stimulus Waveforms

Another of the strategies mentioned above is to use stimulus waveforms that exploit the presence of membrane or synaptic noise. The idea is illustrated in Figure 6, which compares simulations with and

pps	µs/φ	µA
500	30	1000
	500	118
	1000	86
1000	30	1000
	500	118

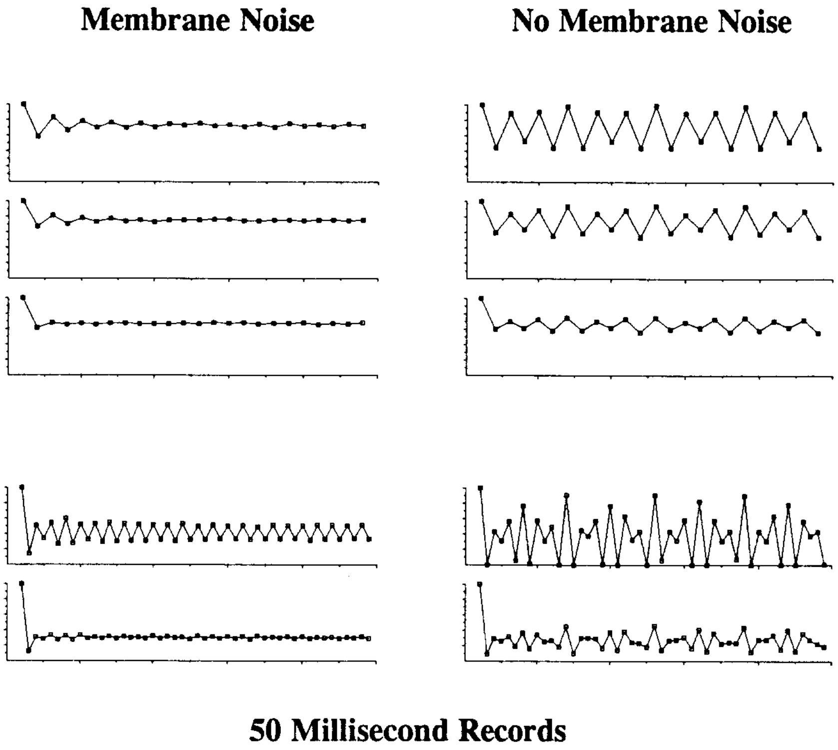

Figure 6. Predicted patterns of neural response to trains of pulses presented at the rates of 500 or 1000 pps. The durations and amplitudes of the pulses for various conditions are indicated in the table on the left side of the figure. Results from simulations using the standard conditions plus membrane noise are shown in the left column and results from simulations using the standard conditions only are shown in the right column.

without membrane noise (level 2.0, see Figure 8 of QPR 7). Stimuli include pulses presented at 500/s and 1000/s, with various pulse durations and amplitudes for each rate. We note that, for the model with membrane noise, the simulation of responses to 30 µs/phase pulses presented at 1000 pps is in excellent agreement with recorded patterns of intracochlear EPs for Ineraid subject SR2 (see Figures 25 and 27 in QPR 7).

Note that neural following of pulses presented at either rate is improved with the addition of membrane noise. For the conditions with membrane noise, magnitudes of neural responses are nearly uniform after pulse 2 for stimulation at 500 pps. The high uniformity observed for 30 µs/phase pulses is slightly improved with increases in pulse duration. Responses for the pulse duration of 1000 µs/phase are essentially identical after pulse 2.

In contrast, following is not particularly good for the condition of stimulation with 30 μs/phase pulses presented at 1000 pps. As described before in QPR 7, an alternating pattern of response is observed over the duration of 50 (illustrated) or 200 ms bursts of pulses. For stimulation at 1000 pps, an increase in pulse duration produces a large improvement in frequency following, as shown in the bottom panel of the left column of Figure 6. Some improvement also is observed for the model simulations without membrane noise, but this improvement is relatively modest. It appears that the presence of membrane noise may be exploited with the use of long-duration low-amplitude pulses.

We plan to evaluate these predictions of the population model in magnitude estimation experiments with implant subjects. We will ask the subjects to report on pitches for a set of stimuli including bursts of pulses with various pulse durations and pulse rates. For example, the set would include 30 μs/phase pulses presented at 1000 pps and 500 μs/phase pulses also presented at 1000 pps. Amplitudes of initial pulses in the bursts will be ramped in a variation of the basic experiment to reduce or eliminate the initial transient in the neural response following pulse 1.

A positive outcome from these experiments, i.e., a correlation of higher perceived pitch with longer pulse duration, would provide the basis for subsequent studies aimed at evaluating the use of long duration pulses in speech processor designs. Obviously, use of such pulses would interact with other desirable attributes in a processor, such as use of relatively high rates of stimulation on single channels and use of nonsimultaneous stimulation across channels. Such tradeoffs could be evaluated in speech reception studies.

Use of High Carrier Rates

As suggested in QPR 7, use of high carrier rates may improve substantially the neural representation of the modulation waveform for SAM pulse trains. This possibility is illustrated further in Figure 7, which shows simulations for three carrier rates and for modulation frequencies of 100, 200, 300 and 400 Hz for each rate. Simulations both with and without membrane noise are shown. The simulations with noise for the 1000 pps carrier are in close agreement with the patterns of intracochlear EPs recorded for Ineraid subject SR2 for the same stimulus conditions (see Figure 31 in QPR 7).

Note that the representations of the 300 and 400 Hz modulation waveforms are improved with increases in carrier rate. For the simulations with membrane noise, the complex pattern of neural response for modulation of a 1000 pps carrier at 300 Hz is replaced by a uniform pattern that better reflects the periodicity of the modulation waveform when the carrier rate is increased to 5000 pps. Also, the presence of two distinct intervals in the neural response to a 1000 pps carrier modulated at 400 Hz is eliminated with an increase in the carrier rate to 5000 pps.

A further increase in the carrier rate, to 10000 pps, produces further improvements in the predicted representation of the modulation waveforms. For example, responses to the cycles of the 400 Hz modulation waveform are essentially identical after the second cycle with the 10000 pps carrier, whereas those responses show an alternating pattern with the 5000 pps carrier.

Improvements also are observed with increases in carrier rate for the simulations without membrane

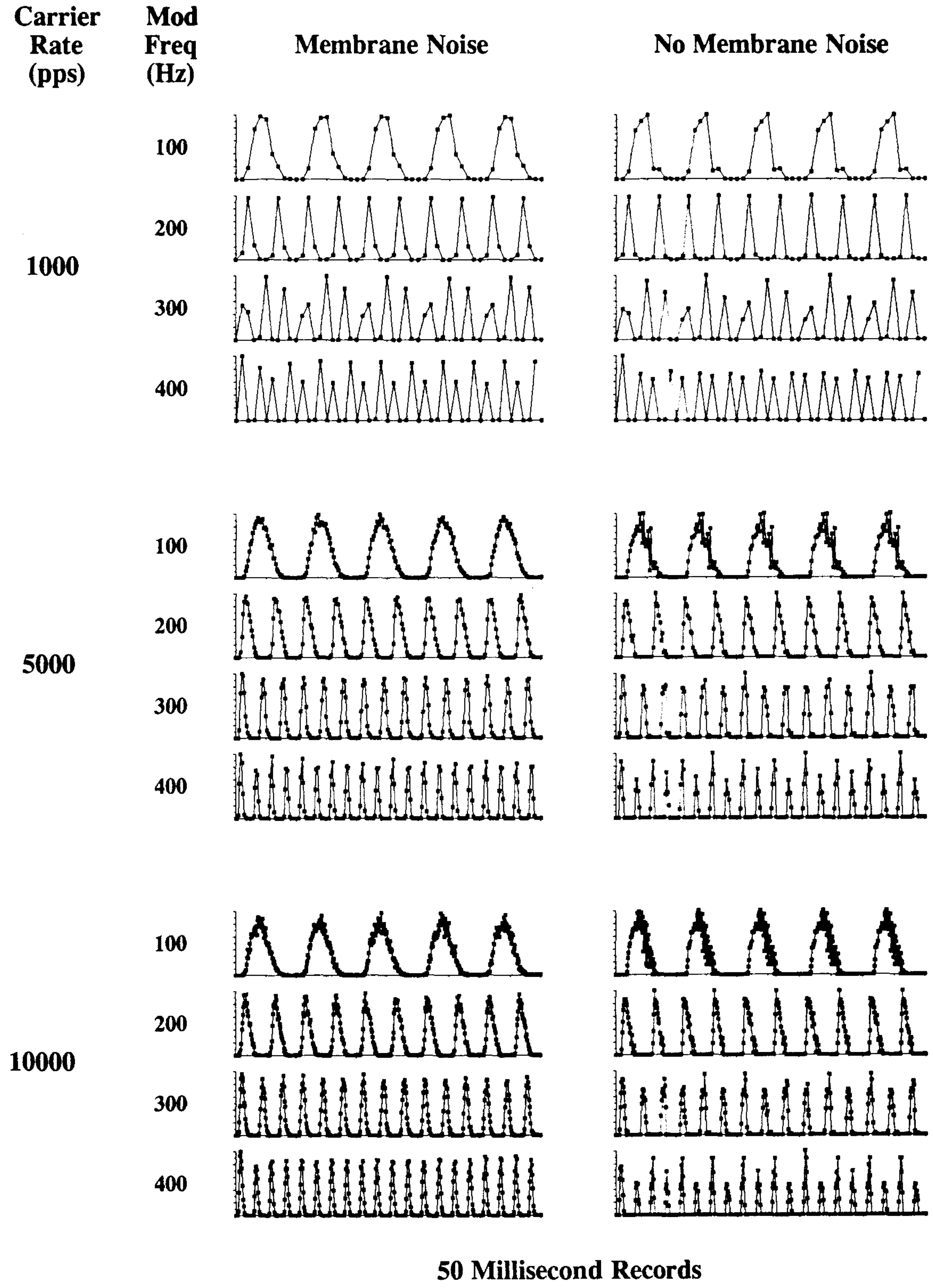

Figure 7. Predicted patterns of neural response to SAM pulse trains, for the indicated carrier rates and modulation frequencies. Results from simulations using the standard conditions plus membrane noise are shown in the left column and results from simulations using the standard conditions only are shown in the right column. The duration of carrier pulses was 30 μs/phase and the depth of modulation was 100 percent.

noise. However, the quality of the representations is higher for the simulations with membrane noise, especially for the 10000 pps carriers. For example, a strong alternating pattern of response is seen for the 400 Hz modulation waveform in the simulation without membrane noise, whereas the pattern is uniform for that modulation waveform in the simulation with membrane noise. The use of relatively high carrier rates may represent an alternative to the use of long duration pulses as a way of exploiting the presence of membrane noise (in this case through jitter in single neuron responses across any of several pulses, from cycle to cycle in the modulation waveform).

We plan to evaluate the use of high carrier rates in CIS processors. Special current sources, with the capability of delivering interleaved pulses at high rates, are under construction for support of these and other studies. (A description of the new current sources will be presented in a future report.)

Additional Strategies

Additional strategies have been identified through modeling studies and through discussions of model results with others. They include:

- Exploiting the presence of membrane noise in another way, by designing stimuli that address neural nodes with the greatest amount of noise, presumably along the narrow-diameter peripheral processes of auditory neurons (see Verveen, 1962; Verveen and Derksen, 1968).
- Introducing a pharmacological agent into the cochlea, e.g., through an osmotic pump (Brown et al., 1993), that would increase the level of membrane noise (this suggestion was offered by Josef Miller in a conversation with Wilson on effects of membrane noise, 1994; also see Junge, 1992, Chapter 12, for a discussion of membrane noise and pharmacological manipulations of the noise).
- Presenting an external noise signal (or signals, on different electrodes) along with the deterministic stimulus, that might enhance effects of existing membrane noise.
- Alternating pulse polarities to address slightly different populations of neurons with successive pulses (this idea was suggested by Bryan Pfingst in a conversation with Wilson, 1994).

Discussion

Recordings of intracochlear evoked potentials have shown that temporal representations with implants are crude and highly limited. Modeling studies, in conjunction with EP recordings, have provided insights into ways those representations might be improved. We expect to evaluate in at least a preliminary way some of these possibilities in the upcoming final three quarters of the current project.

Acknowledgments

A brief description of strategies for the repair of deficits in temporal representations with implants was presented at the *25th Annual Neural Prosthesis Workshop*, held in Bethesda, MD, October 18-21, 1994. The modeling studies described there and in this report were supported by Project IV of NIH Program Project Grant P01-DC00036. Studies with human subjects, including the recordings of intracochlear evoked potentials and psychophysical tests, were supported by the present project. We are grateful to the subjects for their enthusiastic participation and generous contributions of time. We also would like to thank Josef Miller and Bryan Pfingst for their helpful suggestions.

References

Abbas PJ (1993) Electrophysiology. In Tyler RS (Ed.), *Cochlear Implants: Audiological Foundations*. San Diego, CA, Singular Publishing Group, pp 317-355.

Brown JN, Miller JM, Altschuler RA, Nuttall AL (1993) Osmotic pump implant for chronic infusion of drugs into the inner ear. Hear. Res. 70: 167-172.

Junge D (1992) *Nerve and Muscle Excitation*. Sunderland, MA, Sinauer Associates, Inc., Third Edition, pp. 210-226.

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.

Young ED, Sachs MB (1979) Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory-nerve fibers. J. Acoust. Soc. Am. 66: 1381-1403.

III. Plans for the Next Quarter

Our plans for the next quarter include the following:

1. Initial studies with the third patient in the Nucleus percutaneous series (NP-3, January 9-20, 1995) and continued studies with the second patient (NP-1, January 30 through February 10, 1995). Studies with NP-3 will include evaluations of CIS and *spectral peak* (SPEAK) processing strategies. Studies with NP-2 will include repeated measures with the SPEAK strategy, detailed evaluation of CIS processors using more than six channels, and measures of intracochlear evoked potentials.
2. Incorporation of a high-rate, high-resolution A/D converter for the laboratory speech processor system (Burr Brown DEM-PCM1760).
3. Development and initial construction of new current sources, capable of presenting short-duration pulses (< 30 μs/phase) with high fidelity and minimal crosstalk among separate channels.
4. Continued analysis of EP and speech reception data from prior studies.
5. Continued preparation of manuscripts for publication.

Appendix 1

Summary of Reporting Activity for the Period of

August 1 through October 31, 1994

NIH Project N01-DC-2-2401

Reporting activity for the last quarter included three presentations. In addition, two manuscripts were accepted for publication. Citations are listed below.

Presentations

Wilson BS, Finley CC, Zerbi M, Lawson DT: Speech processors for auditory prostheses. Invited presentation, *Neural Prosthesis Workshop*, Bethesda, MD, Oct. 18-21, 1994.

Wilson BS: Cochlear modeling studies. Invited presentation, *Neural Prosthesis Workshop*, Bethesda, MD, Oct. 18-21, 1994. [Work described in this presentation was supported primarily by NIH project P01 DC00036 and secondarily by the present project.]

Loeb GE, Shannon RV, Wilson BS, Zeng F-G, Rebscher S*: 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*. *International Cochlear Implant, Speech and Hearing Symposium*, Melbourne, Australia, Oct. 23-28, 1994. [*Authors are listed in order of a random draw.]

Papers

Wilson BS, Lawson DT, Zerbi M, Finley CC, Wolford RD: New processing strategies in cochlear implantation. Am. J. Otol., accepted for publication.

Wilson BS, Lawson DT, Zerbi M: Advances in coding strategies for cochlear implants. Adv. Otolaryngol. Head Neck Surg., accepted for publication.