

Seventh Quarterly Progress Report

February 1 through April 30, 1997

NIH Project N01-DC-5-2103

Speech Processors for Auditory Prostheses

Prepared by

Blake Wilson, Charles Finley, Marian Zerbi, Dewey Lawson and Chris van den Honert

Center for Auditory Prosthesis Research

Research Triangle Institute

Research Triangle Park, NC 27709

Contents

[I. Introduction](#)

[II. High Rate Studies, Subject SR2](#)

[III. Plans for the Next Quarter](#)

[IV. Acknowledgments](#)

[Appendix 1: Summary of Reporting Activity for this Quarter](#)

I. Introduction

One of the principal objectives of this project is to design, develop, and evaluate speech processors for implantable auditory prostheses. Ideally, the processors will represent the information content of speech in a way that can be perceived by implant patients. Another principal objective is to develop new test materials for the evaluation of speech processors, given the growing number of cochlear implant subjects enjoying levels of performance too high to be sensitively measured by existing tests.

Work in the present quarter included:

- Discussion with consultant Sigfrid Soli to review and refine plans for upcoming studies with recipients of bilateral implants. The discussion included considerations for the design of processing strategies for coordinated stimulation of the two sides and considerations for the design and use of tests to evaluate possible advantages of such stimulation in noisy environments. Dr. Soli graciously agreed to provide a test system developed at HEI, and to process special German language test materials for use in that system. Dr. Soli also will participate in speech reception studies planned for this fall with a German-speaking subject having full insertions of the Med El electrode array on both sides. Dr. Soli visited RTI on March 5 and 6.
- Continued preparation for studies with bilateral patients, including (a) development and installation of test instruments; (b) initial development, with the cooperation and assistance of the University of Innsbruck and Med El, of an interface system to allow full and simultaneous laboratory control of two Med El implanted receivers; (c) development of processing strategies for coordinated stimulation of the two sides; (d) further development of psychophysical tests to measure sensitivities to localization cues such as interaural timing and amplitude differences; and (e) initial development of DSP code for implementation of various processing strategies.

- Final studies with Nucleus percutaneous subject NP2, to evaluate possible learning effects with chronic use of a CIS processor. The Geneva/MEEI/RTI processor was fit on January 31 and tested again on February 25, after approximately one month of daily use.
- Final studies with Nucleus percutaneous subject NP4, to measure spatial patterns of stimulation and neural response with recordings of intracochlear EPs (April 7-8).
- Studies with Ineraid subject SR10, including (a) measures of intracochlear evoked potentials (EPs) for unmodulated and sinusoidally amplitude modulated pulse trains with pulse rates up to 10000/s; (b) measures of intracochlear EPs for pairs of identical pulses with several amplitudes and a wide range of interpulse intervals; (c) measures to verify operation of and identify remaining problems with new stimulating and recording systems for EP studies; (d) measures of electrically evoked auditory brainstem responses (EABRs) for all six electrodes; (e) measures of consonant identification for processors with different numbers of electrodes; (f) measures of consonant identification for several single-channel processors using different electrodes; and (g) measures of consonant identification for six-channel processors using various rates of stimulation up to 10000 pulses/s/channel. SR10 visited RTI for the period of April 21 to May 2.
- Ongoing studies with Ineraid subject SR2 (usually one morning per week). Studies in this quarter included measures of intracochlear EPs for a variety of stimuli with high pulse rates, measures of intracochlear EPs for pairs of identical pulses with several pulse amplitudes and a wide range of interpulse intervals, measures to verify operation of and identify remaining problems with new stimulating and recording systems for EP studies, and preliminary evaluations of speech processor designs using relatively high rates of stimulation.
- Continued development of the speech reception and evoked potential laboratories
- Presentation of project results in two invited lectures.
- Preparation for four invited lectures to be presented in May, one at the Vth International Cochlear Implant Conference in New York City, May 1-2, two at the 131st annual meeting of the American Otological Society in Scottsdale, AZ, May 10-11, and one for the Department of Biomedical Engineering at Johns Hopkins University, May 30.
- Completion of two manuscripts submitted in conjunction with the AOS meeting (submitted April 25).
- Continuing preparation of other manuscripts, including an invited editorial on "The future of cochlear implants," for the British Journal of Audiology.
- Continued analysis of speech reception and evoked potential data from prior studies.

In this report we present initial results from electrophysiological and psychophysical studies involving rates of stimulation up to 10000 pulses/s. Such results provide basic information for the design and application of speech processors using relatively high rates. Results from other studies and activities indicated above will be presented in future reports.

II. High Rate Studies, Subject SR2

Recordings of intracochlear evoked potentials (EPs) have allowed us to visualize temporal patterns of neural responses for a wide variety of electrical stimuli. Subjects for such recordings have been patients with percutaneous access to their implanted electrodes, i.e., users of the Ineraid device or a research version of the Nucleus device. Stimuli to date have included (a) trains of identical pulses with various pulse rates, pulse amplitudes, and burst durations; (b) pairs of identical pulses with a wide range of interpulse intervals; (c) pairs of pulses with several fixed interpulse intervals and various amplitudes for the first pulse; (d) sinusoidally amplitude modulated (SAM) pulse trains with various carrier rates, modulation frequencies, modulation depths, and burst durations; (e) the pulsatile outputs of a single-channel speech processor; and (f) a single probe pulse following a masker consisting of a pulse train or a SAM pulse train. Use of a subtraction technique has allowed us to investigate responses to stimuli with pulse rates in excess of 1000/s, which otherwise would be difficult to interpret because of overlapping among successive evoked potential waveforms.

Studies also have been conducted to evaluate possible psychophysical correlates of the recorded patterns of neural response for various stimuli. For example, we have evaluated scaling of modulation frequencies for SAM pulse trains with a relatively wide range of carrier rates.

In this report we will present recordings for one subject of neural population responses to trains of identical pulses and SAM pulse trains, for rates of stimulation between 100 and 10000 pulses/s. We also will present preliminary results from psychophysical experiments in which the subject was asked to scale pulse rate for unmodulated pulse trains, and modulation frequency for SAM pulse trains, according to perceived pitch. Rates for the unmodulated pulse trains varied

between 100 and 600 pulses/s, and modulation frequencies for SAM pulse trains varied between 100 and 600 Hz. Carrier rates for the SAM pulse trains varied between 500 and 10000 pulses/s.

Subject SR2 participated in these initial studies. We have collected but not yet fully analyzed "high rate" data for subjects SR10 and NP2. We also plan to collect data for subjects SR3, SR9 and SR16 during the next quarter. Subjects SR9, SR10 and SR16 have relatively low levels of speech reception performance with their implants, while SR2 enjoys a very high level of performance. Performance for NP2 is good but not as high as that for SR2. Analysis of the data for all subjects will be completed by the end of the quarter, and results across subjects will be presented in our next progress report.

System used for recordings of intracochlear evoked potentials

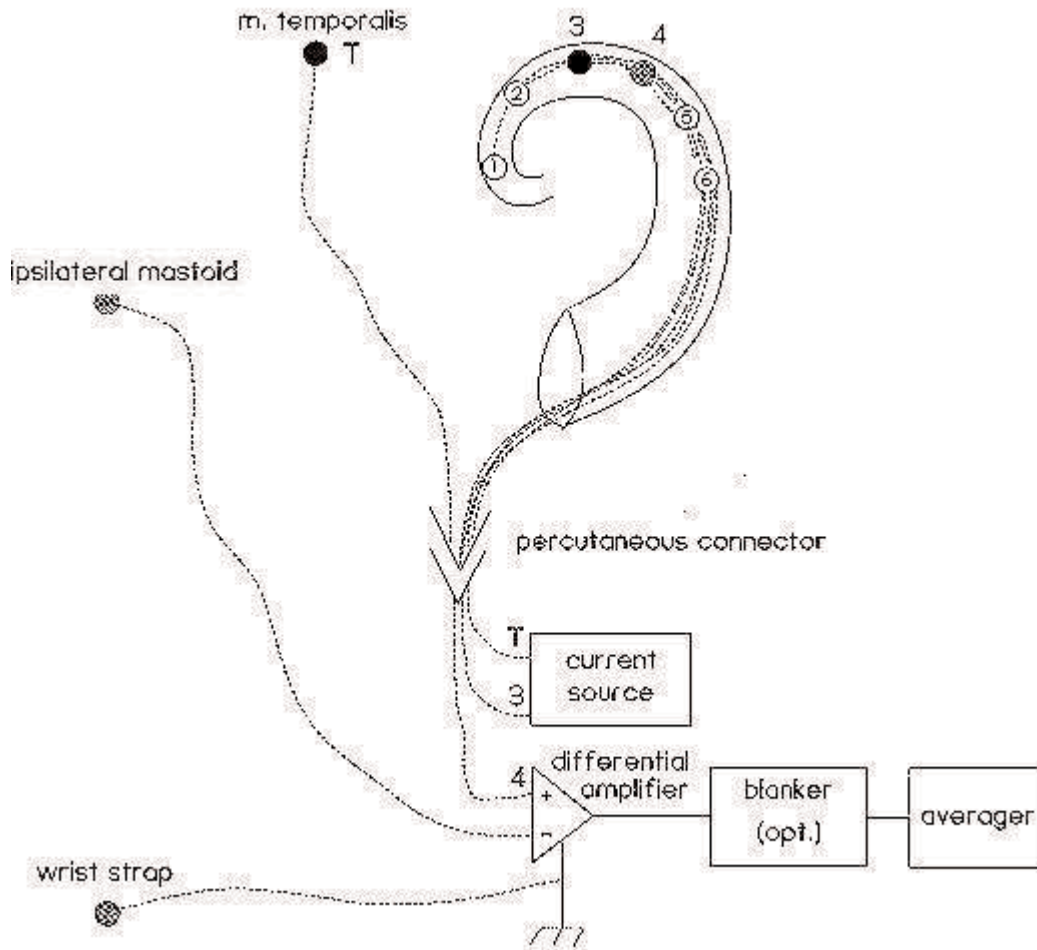


Fig. 1. Apparatus for recording intracochlear evoked potentials.

The system we use for recordings of intracochlear EPs is illustrated in Fig. 1 (Wilson *et al.*, 1994; Wilson *et al.*, in press). Intracochlear potentials are measured differentially between an unstimulated electrode in the implant and an electrode at the ipsilateral mastoid. Body potential is measured with a reference electrode at the wrist. Stimuli are delivered between an intracochlear electrode and a reference electrode implanted in the temporalis muscle (monopolar stimulation) or between two intracochlear electrodes (bipolar stimulation). A fast recovery amplifier is used to restore sensitivity of recording as soon as possible after saturation of the input by stimulus pulses. In addition, an equal number of sweeps for negative leading and positive leading biphasic pulses are summed to cancel components of the artifact. With these techniques, the blinker circuit generally is not necessary for clear separation of EPs from residual artifacts, and was not used in obtaining the data shown in this report.

This arrangement for recording intracochlear EPs can be used with implant systems having direct percutaneous access to the implanted electrodes, for instance the Ineraid electrode and percutaneous connector as shown in Fig. 1. We also have made recordings with subjects implanted with a percutaneous connector version of the 22-electrode Nucleus

array.

Examples of recordings for rates of stimulation at and below 1000 pulses/s

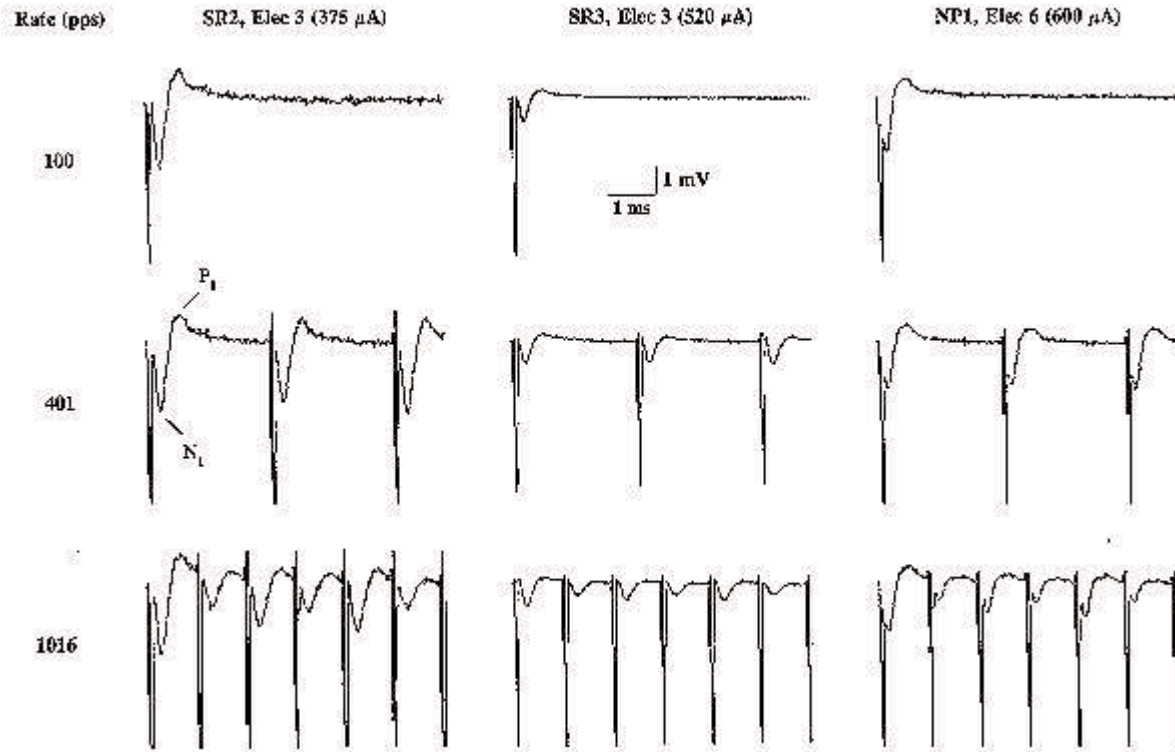


Fig. 2. Intracochlear evoked potentials for subjects SR2, SR3 and NP1.

Examples of recordings for three subjects, including subject SR2, are presented in Fig. 2. The stimuli in these examples were 200 ms trains of identical 33 μ s/phase pulses, with pulse rates of 100, 401 and 1016/s. Monopolar stimulation was used. For each of the three subjects, the amplitude of the pulses was adjusted to produce a most-comfortable-loudness (MCL) percept for the 1016 pulses/s condition. This amplitude was held constant for the lower rates, producing percepts with lower loudnesses.

The figure shows the first 6 ms of the 200 ms records for each subject. The large downward "spikes" are residual (uncanceled) artifact during and shortly after presentations of stimulus pulses. Following these pulse artifacts are neural evoked potentials, with a negative peak (N1) approximately 250 μ s after pulse onset and a positive peak (P1) approximately 600 μ s after pulse onset. The magnitudes of the EPs, as measured by the absolute difference between the peak voltages at N1 and P1, range up to several millivolts for the subjects and conditions of Fig. 2.

For low rates of stimulation EPs reflect the identical amplitudes of the pulses. For the 401 pulses/s conditions in Fig. 2, for example, nearly equal EPs are observed across the three illustrated pulses for each subject. At higher rates an alternating pattern of response is observed, with a large EP following the first pulse, a much diminished EP following the second pulse, a partially recovered EP following the third pulse, another small EP following the fourth pulse, and so on. Such patterns may in part be an expression of the refractory properties of auditory neurons. Presumably, many neurons are available for stimulation by the first pulse, when the nerve is at rest and the excitability of each neuron is at its maximum. Neurons stimulated by that pulse then become refractory to subsequent stimulation. At the time of the second pulse, approximately 1 ms later, those neurons would be in a period of relative refraction (Hartmann *et al.*, 1984; Parkins, 1989), with reduced excitability. Thus, not as many neurons would be expected to respond to the second pulse. At the time of the third pulse, neurons stimulated by the first pulse but not the second will have recovered much (but not all) of their initial excitability. More neurons might be expected to respond to the third pulse than to the second. The pattern of alternation between relatively large and relatively small EPs can persist for hundreds of milliseconds at particular rates for a given subject and stimulating electrode (Wilson *et al.*, 1995).

The magnitudes of EPs and patterns of response across rate of stimulation vary widely across subjects and often across electrodes within subjects (Wilson *et al.*, in press). These differences may reflect differences in refractory properties of the stimulated neurons, the number of neurons participating in the response, subthreshold integration of sequential pulses at neural membranes, and other factors (see Wilson *et al.*, 1994).

9. Subtraction technique for pulse rates higher than 1000/s

Analysis of recordings for pulse rates above about 1000/s is complicated by the fact that EPs for successive pulses overlap when the interval between the pulses is shorter than approximately 1 ms. Thus, the waveform following a particular pulse in a train of pulses reflects not only the response to that pulse but also the trailing parts of the EPs to the the prior pulse or pulses.

A way to derive the response to the Nth pulse in a recording containing overlapping EPs is illustrated in Fig 3. The effects of all prior stimuli and responses overlapping with the Nth response can be removed by subtracting a record for a burst with N-1 pulses, as shown in the right column of the figure. This method is time consuming, requiring N recordings for a train with N pulses, but it allows us to study responses to individual pulses for pulse rates exceeding 1000/s.

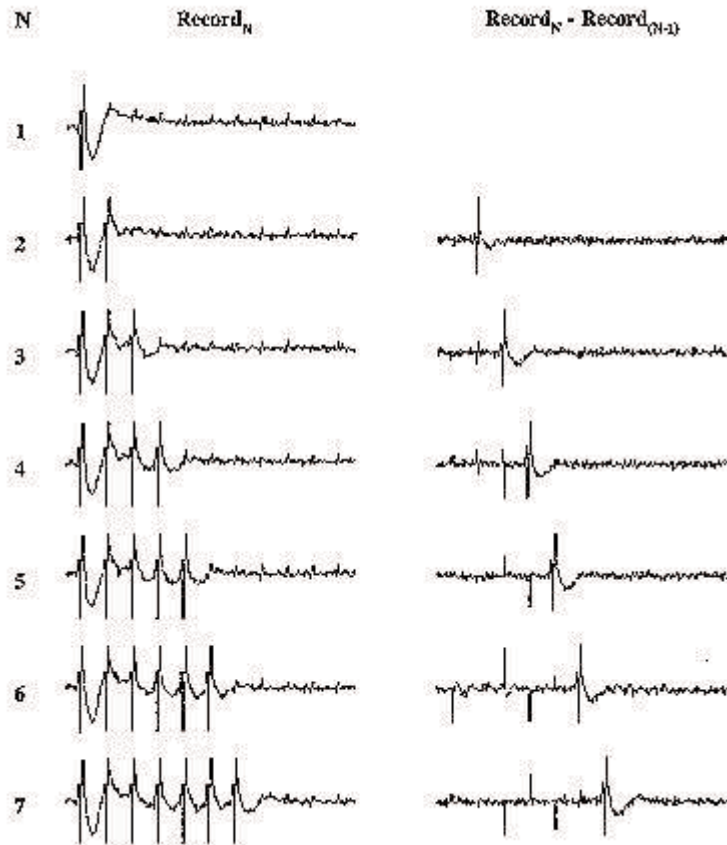


Fig. 3. Example of subtraction technique used for measurement of evoked potentials to pulses presented at high rates. The three columns show the number of pulses (N), the record of the pulse artifact(s) and response(s) collected with that number of pulses (Record_N), and the derived EP obtained by subtracting a record collected with N-1 pulses from a record collected with N pulses (Record_N - Record_(N-1)). Pulses were presented at the rate of 2033/s for conditions involving more than one pulse. Data are from studies with subject SR2.

Neural representations of unmodulated pulse trains

Results from studies with subject SR2 using both low and high rates of stimulation are presented in Fig. 4. Patterns of responses for low rates of stimulation are presented in the left column and patterns of responses for high rates in the right column. The figure shows the magnitudes of the EPs following each stimulus pulse. The magnitudes were

measured as the absolute difference between the voltages at the N1 and P1 peaks in each EP waveform. These magnitudes then were normalized to the magnitude of the EP following the first pulse for each condition. The subtraction technique was used to derive EP magnitudes for the conditions in the right column. Stimulus pulses were delivered to electrode 3 in the subject's Ineraid implant, with reference to a remote electrode in the temporalis muscle, and intracochlear voltages were recorded with the adjacent electrode 4 in the implant, with reference to an electrode at the ipsilateral mastoid. Note that different time scales are used for the left and right columns of Fig. 4.

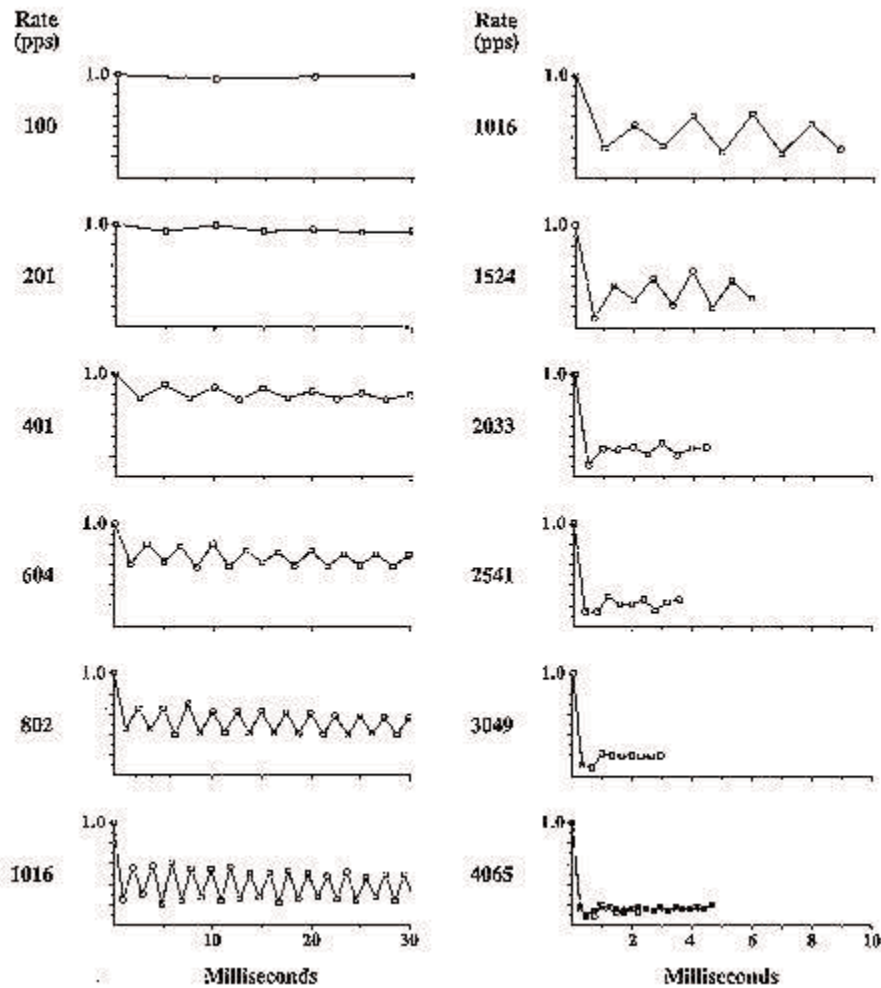


Fig. 4. Magnitudes of evoked potentials for stimulation of subject SR2's intracochlear electrode 3 and recording with intracochlear electrode 4. Patterns of EP magnitudes for relatively low pulse rates are shown in the left column and those for relatively high rates in the right column. Responses for the conditions in the right column were derived using the subtraction technique illustrated in Fig. 3. The EP magnitudes are normalized to the magnitude of the EP following the first pulse for each condition. The pulse amplitude used across all conditions was $375 \mu\text{A}$, and the pulse duration was $33 \mu\text{s}/\text{phase}$. The filled symbols in the panel for 4065 pulses/s stimulation show results from a repeated measure collected during a separate visit to the laboratory by SR2. Note that different time scales are used for the left and right columns. (The first, third, and sixth panels in the left column represent the same conditions as those for the recordings shown in the left column of Fig. 2.)

The figure shows that EPs for this subject and stimulating electrode reflect the identical amplitudes of the stimulus pulses with nearly identical magnitudes of response for the pulse rates of 100 and 201/s. At 401 pulses/s an alternating pattern of responses is observed, with relatively large responses for odd-numbered pulses and relatively small responses for even-numbered pulses. Alternating patterns of responses also are observed for the higher rates, with progressively greater decrements in the response to pulse 2 and progressively greater differences between responses for odd- and even-numbered pulses as rate is increased up to 1016 pulses/s. At somewhat higher rates, the alternating pattern is replaced by more complicated patterns of responses, e.g., at 1524 pulses/s. At still higher rates, results for SR2 show a return to a simpler pattern of response, with uniform magnitudes of sequential EPs for identical stimulus pulses after the first millisecond of stimulation. This is most evident in the patterns of responses for the 3049 and 4065 pulses/s

conditions.

Modeling studies suggest that such uniform patterns beyond the first millisecond may be a result of a more stochastic response among neurons to the stimulus pulses (see, e.g., Fig. 8 and the accompanying discussion in Wilson *et al.*, 1994). When pulses are presented at high rates, low levels of neural membrane noise at nodes of Ranvier may interact with the pulses to produce stochastic independence among neurons. Slight variations in neural threshold due to membrane noise may introduce a "jitter" in firing times across neurons for rapidly presented pulses. The effect of such jitter would be expected to increase with time after the beginning of a train of pulses, as initially small differences in discharge histories among neurons increased. After a relatively short period, differences in discharge histories may produce a high level of stochastic independence among neurons.

Such a mechanism would allow different subpopulations of neurons in the excitation field to respond to sequential pulses (Parnas, 1996; Wilson *et al.*, 1994). The total number of neurons responding to any one pulse would in general be small compared to the number stimulated by single pulses at low rates of stimulation. Thus, one might expect relatively small EPs for high rates in conjunction with relatively uniform EPs from pulse to pulse for trains of identical pulses.

Responses during the initial millisecond at relatively high rates may reflect a response of many neurons to the first pulse, when excitability of the neurons is at its maximum, followed by a subsequent depression in response(s) when those same neurons are in a period of absolute refraction or early in a period of relative refraction. Responses during the initial millisecond also may reflect effects of temporal integration at neural membranes. In particular, note that responses to pulse 2 increase as the rate is increased from 1524 to 4065 pulses/s. For some neurons, stimulation with pulse 1 alone might not be sufficient for excitation. Integration of pulses 1 and 2, however, might be sufficient. In such cases, a neuron would respond to pulse 2.

The possibility of temporal integration effects is illustrated further in Fig. 5, which shows patterns of response for rates up to 10162 pulses/s. For rates at and above 5039 pulses/s, magnitudes of response for the second and third pulses are greater than the magnitudes of response to subsequent pulses in the "well" between 0 and 1 ms. These results suggest that, for the conditions of Fig. 5, the population of neurons not stimulated by pulse 1 may be relatively small and that temporal integration of subthreshold pulses does not occur when the pulses are separated by more than about 0.5 ms.

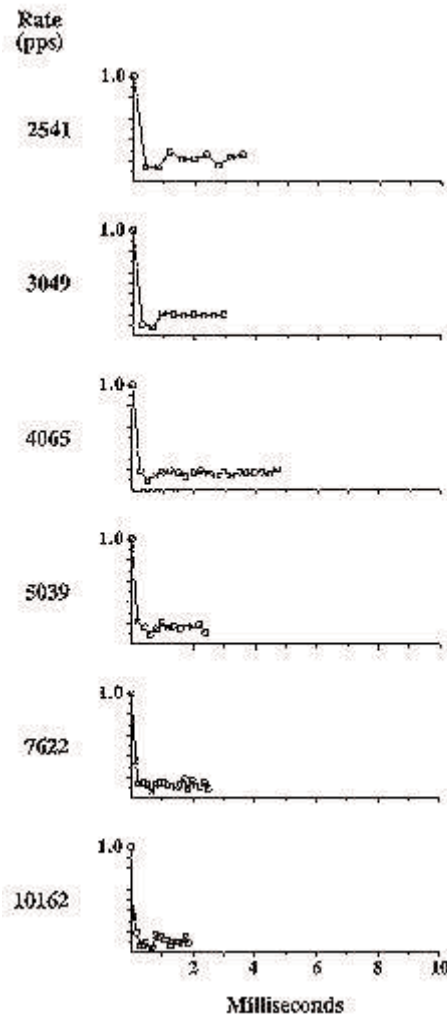


Fig. 5. Magnitudes of evoked potentials, as in Fig. 4, but here for higher rates of stimulation. The panels for the 2541, 3049, and 4065 pulses/s conditions are the same as those in Fig. 4, and are repeated here to illustrate increases in the response to pulse 2 as rate is increased from 2541 to 10162 pulses/s.

The alternating and more complex patterns of responses observed for SR2 at rates between about 400 and 2500 pulses/s may indicate limitations on the transmission of stimulus information to the central nervous system. That is, such patterns appear to reflect the properties of the auditory nerve as well as properties of the stimulus. As described in prior reports (e.g., Wilson *et al.*, 1994; 1995; in press), and as indicated in Fig. 2, alternating and complex patterns also have been observed for other subjects at rates between 400 pulses/s and the maximum studied rate of 1000 pulses/s.

Neural representations of modulated pulse trains

A majority of the processing strategies in current clinical use, including the SPEAK and CIS strategies, employ modulated pulse trains as stimuli. To obtain information on how such stimuli are represented in the population responses of the auditory nerve, we have recorded intracochlear EPs for sinusoidally amplitude modulated (SAM) pulse trains with seven subjects. Studies with most of these subjects included wide ranges of modulation frequencies and various carrier rates up to 1016 pulses/s. Studies with SR2 have included the carrier rate of 4065 pulses/s.

A representative set of results for carrier rates at and below approximately 1000 pulses/s is presented in Fig. 6 (exact carrier rates and modulation frequencies are indicated in the caption to Fig. 6). The stimuli in this case were delivered to electrode 3 in subject SR3's Ineraid implant, with reference to a remote electrode in the temporalis muscle, and neural responses were recorded with electrode 4 in the implant, with reference to an electrode at the ipsilateral mastoid. The carrier level was adjusted to produce a most comfortable loudness (MCL) percept for 400 Hz modulation of a 1000 pulses/s carrier (bottom right condition in Fig. 6). This level was held constant across all conditions, producing somewhat lower loudnesses for the remaining modulation frequencies for the 1000 pulses/s carrier and for all

conditions with the 500 and 250 pulses/s carriers.

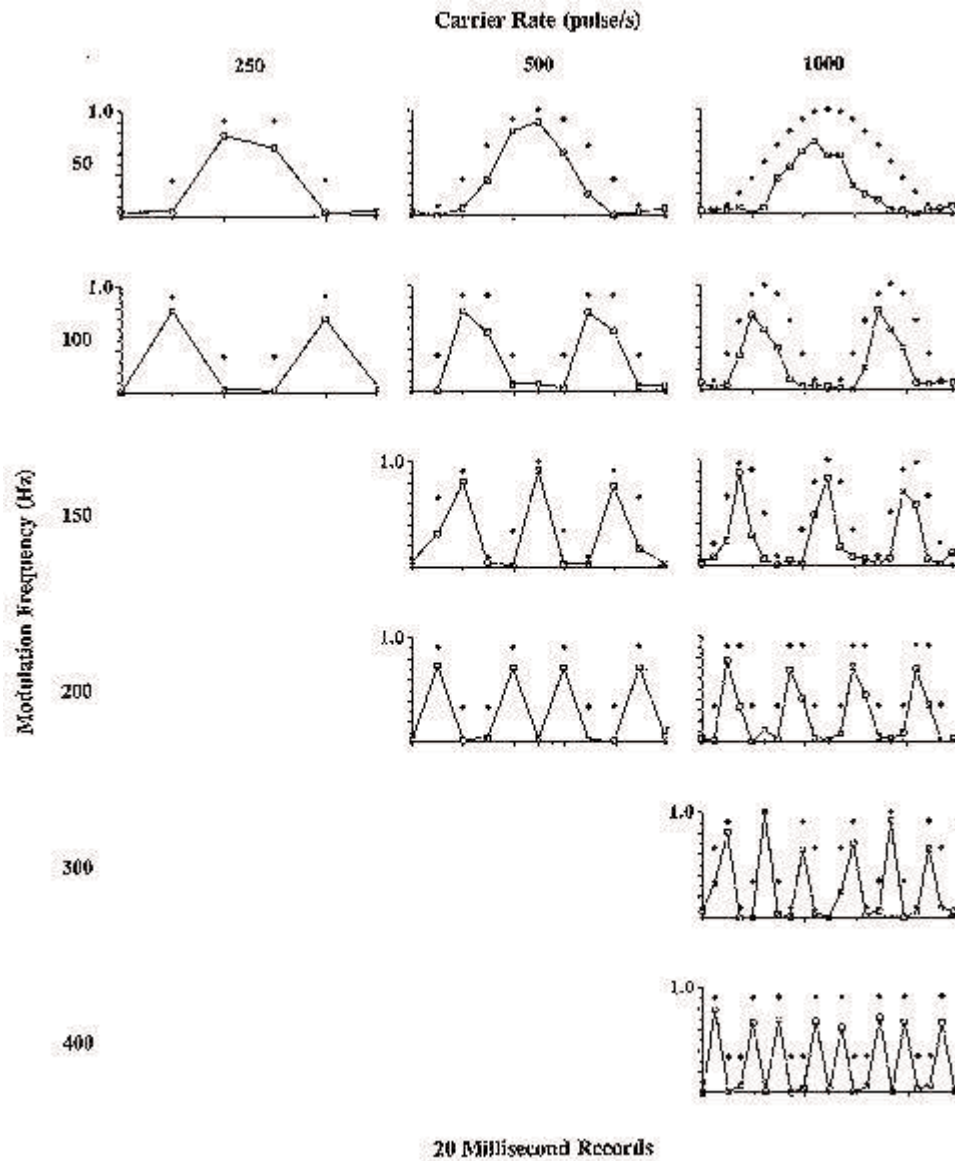


Fig. 6. Pulse amplitudes (filled diamonds) and evoked potential magnitudes (connected open squares) for sinusoidally amplitude modulated (SAM) pulse trains. Normalized values are shown for both measures, with a value of 1.0 corresponding to the maximum pulse amplitude or maximum EP magnitude across all conditions. The stimuli were generated by modulating 250, 500 and 1000 pulses/s carriers at the indicated frequencies. The carrier rates were adjusted slightly to achieve uniform intervals between pulses with a 16.4 μ s sampling interval. The adjusted rates were 251, 504 and 1016 pulses/s. The adjustments also scaled the modulation frequencies; the frequencies for the 251 pulses carrier remained at 50 and 100 Hz, whereas the frequencies for the 504 pulses/s carrier were 50, 101, 151 and 202 Hz, and the frequencies for the 1016 pulses/s carrier were 51, 102, 152, 203, 305 and 407 Hz. Data are from studies with subject SR3. The carrier level for all conditions was 600 μ A, and the pulse duration was 33 μ s/phase. Stimuli were delivered to intracochlear electrode 3 and recordings of neural responses were made with intracochlear electrode 4.

The amplitudes of the stimulus pulses for each condition are indicated by solid diamonds and the magnitudes of the EPs following each pulse by connected open squares. The pulse amplitudes are normalized to the maximum amplitude across all conditions, and the EP magnitudes are normalized to the maximum magnitude across all conditions. The peak amplitude was 600 μ A and the pulse duration was 33 μ s/phase. While only the first 20 ms are shown in Fig. 6, the duration of each SAM pulse train was 200 ms.

The patterns of responses appear to reflect both sampling of the modulation waveform by the carrier pulses and the nonlinear properties of auditory neurons. Examples of apparent refractory effects may be seen in the panels for 100 Hz

modulation of 500 and 1000 pulses/s carriers. The third and fourth pulses for the 500 pulses/s carrier have identical amplitudes and yet the neural response to the fourth pulse is substantially lower. For the 1000 pulses/s carrier the sixth pulse is higher in amplitude than the fifth pulse and yet the neural response to the second of those two pulses is again lower in magnitude.

Note also that when the modulation frequency is low compared to the carrier rate the pattern of EP magnitudes approximates the pattern of pulse amplitudes. For 50 Hz modulation of the 500 pulses/s carrier, for example, the pattern of neural responses looks almost sinusoidal, with a somewhat closer approximation to the stimulus pulses in the first half of the modulation cycle. As the modulation frequency is increased, the asymmetry of responses in each modulation cycle increases. For 100 Hz modulation of the 500 pulses/s carrier, for example, a "peaking" of the responses is observed in the first half of each modulation cycle.

Further increases in modulation frequency produce more complex patterns of responses. The pattern of responses for 150 Hz modulation of the 500 pulses/s carrier reflects the overall frequency of modulation but also shows large variations from cycle to cycle. The "sampling" of the sinusoidal modulation waveform becomes progressively more sparse with increases in modulation frequency. The sparse sampling for the 150 Hz modulation condition only crudely reflects the modulation waveform. As the modulation frequency approaches one half the carrier rate (the "Nyquist frequency," see Rabiner and Schafer, 1978), multiple intervals and other anomalies can appear in the stimuli and in the patterns of responses. Multiple intervals appear, for example, in the stimuli and pattern of responses for 200 Hz modulation of the 500 pulses/s carrier. The time between peaks in the response alternates between long (6 ms) and short (4 ms) intervals. Neither of these intervals corresponds to the period of the modulation waveform (5 ms).

The effects just described for the 500 pulses/s conditions scale with carrier rate. For the 1000 pulses/s carrier, for example, a highly complex pattern of responses is observed at the modulation frequency of 300 Hz, and a pattern of responses with two distinct intervals is observed at the modulation frequency of 400 Hz. For the 250 pulses/s carrier, two distinct intervals are observed in the pattern of responses at the modulation frequency of 100 Hz, although the 20 ms segment of the record presented in Fig. 6 is too short to show both intervals for this particular condition.

Percepts reported by subjects SR2 and SR3 when listening to these stimuli are consistent with the recorded patterns of responses. For the 500 pulses/s carrier conditions, the subjects report increases in pitch with increases in modulation frequency. The percepts elicited with relatively low modulation frequencies are described as smooth and tonal. However, the percept for the 150 Hz modulation condition is described as sounding rough and complex. Also, the percept for the 200 Hz modulation condition is described as combining at least two separate tones. For the 1000 pulses/s carriers, the percepts for the 150 and 200 Hz modulation conditions are described as relatively smooth and tonal, particularly the percept for the 200 Hz modulation condition. The percepts for the two lower modulation frequencies also are described as smooth and tonal, as before. For the higher modulation frequencies, however, a rough and complex percept is again reported, but this time at the modulation frequency of 300 Hz, and a multitonal percept is again reported, but this time at the modulation frequency of 400 Hz. These reports for the higher carrier rate are consistent with the scaling of neural response patterns with changes in carrier rate, as described above.

In broad terms, the results of Fig. 6 suggest that the carrier rate in CIS and other processors should be 4 to 5 times the highest frequency in the modulation waveforms for a smooth and unambiguous representation of those waveforms. Busby *et al.* (1993) have offered this same suggestion, based on results from their psychophysical studies with patients using the Nucleus device.

Additional studies have been conducted by us to evaluate effects of even higher carrier rates on neural representations of modulation waveforms (Wilson *et al.*, 1996). Results for SR2 are presented in Fig. 7. The stimuli were delivered to intracochlear electrode 3, with reference to a remote electrode in the temporalis muscle, and the neural responses were recorded with intracochlear electrode 4, with reference to an electrode at the ipsilateral mastoid. The carrier level across all conditions was 475 μ A. The pulse duration was 33 μ s/phase.

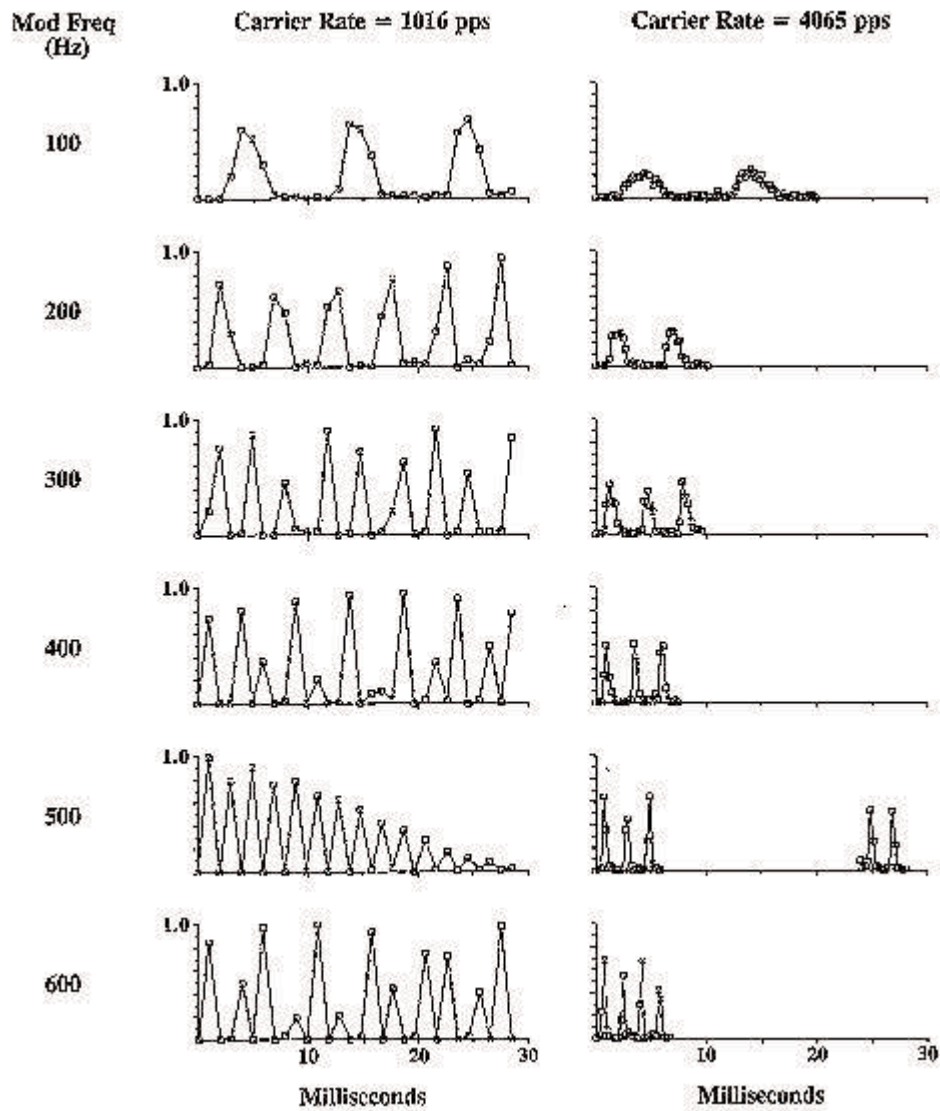


Fig. 7. Evoked potential magnitudes for SAM pulse trains with the carrier rates of 1016 and 4065 pulses/s. EP magnitudes are normalized to the maximum value across all conditions. The modulation frequencies used in the studies of this figure are somewhat different from those used in the studies of Fig. 6 and are exact for all conditions. Responses for the high rate carrier were derived using the subtraction technique illustrated in Fig. 3. Data are from studies with subject SR2. The carrier level for all conditions was 475 μ A, and the pulse duration was 33 μ s/phase. Stimuli were delivered to intracochlear electrode 3 and recordings of neural responses were made with intracochlear electrode 4.

The comparison in Fig. 7 is between carrier rates of 1016 pulses/s and 4065 pulses/s, for the modulation frequencies of 100, 200, 300, 400, 500 and 600 Hz. Note that the combinations of carrier rate and modulation frequencies are somewhat different from those of the corresponding conditions in Fig. 6 (see caption to Fig. 6). Note also that 30 ms records are presented in Fig. 7, rather than the 20 ms records found in Fig. 6. The EP magnitudes in Fig. 7 are normalized to the maximum magnitude across all conditions.

Results for the 1016 pulses/s carrier show simple representations of the modulation frequency for the 100 and 200 Hz modulation conditions. The pattern of responses becomes more complicated at a modulation frequency of 300 Hz. At 400 Hz the pattern is both complicated and no longer reflects the period of the modulation waveform. The first interval between major peaks in the response (between pulses 2 and 5) roughly approximates the period, but subsequent intervals are much longer than the period. (The difference between this pattern of responses and the pattern for the corresponding condition in Fig. 6 is consistent with the slight difference in the combinations of modulation frequency and carrier rate.)

Close approximation of the modulation frequency to the Nyquist frequency, as with the 500 Hz modulation condition, produces a pattern of stimulation in which pulses of relatively high amplitudes alternate with pulses of relatively low

amplitudes. The difference between high and low amplitude pulses wanes and waxes at a "beat frequency" equal to the difference between the modulation and Nyquist frequencies, in this case 8 Hz. Modulation precisely at the Nyquist frequency would produce a series of alternating pulses with fixed amplitudes at two levels, the levels depending on the fixed phase offset between the pulses and the modulation waveform.

The pattern of responses in Fig. 7 for 500 Hz modulation of the 1016 pulses/s carrier reflects the pattern of stimulation, i.e., alternating high and low EP magnitudes with the difference in high and low magnitudes diminishing over the half-period of the 8 Hz beat frequency. In addition, the response to pulse 4 is low compared to the response to pulse 2, and the response to pulse 8 is low compared to the response to pulse 6. This latter pattern probably reflects refractory properties of the neurons, as discussed before in connection with Figs. 2 and 6.

When the modulation frequency exceeds the Nyquist frequency a phenomenon called "aliasing" occurs, in which the pattern of stimulation for a given modulation frequency above the Nyquist frequency is identical (except for a possible phase offset) to the pattern that would be obtained for a modulation frequency below the Nyquist frequency by an equal amount. For example, with a 1000 pulses/s carrier identical patterns of stimulation would be produced with modulation frequencies of 400 and 600 Hz (the pattern resulting from aliasing by the 600 Hz modulation is like the uncorrupted pattern produced by 400 Hz modulation).

For the present conditions the 400 and 600 Hz modulation frequencies are not quite equally distant from the Nyquist frequency (508 Hz). However, they are close enough to produce similarities in the patterns of stimulation and responses. A complicated pattern of response is again observed for the 600 Hz modulation condition, that does not reflect the frequency of the modulation waveform.

Three regions of responses can be identified for the 1016 pulses/s carrier. At relatively low modulation frequencies, i.e., 100 and 200 Hz, the responses simply represent the modulation waveform. At somewhat higher modulation frequencies complex patterns of response are observed. Those patterns do not correspond to any details of the modulation waveform and, indeed, at the modulation frequency of 400 Hz do not represent the modulation frequency. Severe sampling artifacts occur as the Nyquist frequency is approximated or exceeded by the modulation frequency. Results of such artifacts can be seen in the patterns of responses for the 500 and 600 Hz modulation conditions.

Patterns of responses for the 4065 pulses/s carrier show simple representations for all modulation frequencies included in the studies of Fig. 7 (responses for the high rate carrier were derived using the subtraction technique of Fig. 3). The patterns of responses follow closely the patterns of stimulation for the modulation frequencies of 400 Hz and lower. The distortions noted before for 300 and 400 Hz modulation of the 1016 pulses/s carrier are eliminated with an increase in carrier rate to 4065 pulses/s. At the higher modulation frequencies of 500 and 600 Hz, the patterns of responses for the 4065 pulses/s carrier show a shallow alternation between high and low peaks for successive cycles of the modulation waveform. For the 500 Hz condition this alternation may be damped or absent after the initial cycles, as suggested by the pattern of responses for two cycles beginning about 24 ms after the onset of the burst.

Additional aspects of the responses for the 4065 pulses/s carrier are that (a) the peak magnitudes are lower than those for the 1016 pulses/s carrier and (b) the responses from pulse to pulse are smooth and continuous within modulation cycles. These aspects are consistent with the idea that high rate stimuli elicit a more stochastic pattern of responses within and among neurons than low rate stimuli, as described above in connection with responses to unmodulated pulses presented at high rates. However, the gradual increase in pulse amplitudes at the beginning of the burst for SAM pulse trains may reduce or eliminate the initial "transient" response observed with unmodulated pulse trains.

Perception of modulated pulse trains

Findings from recordings of intracochlear EPs led us to measure effects of carrier rate on psychophysical scaling of modulation frequencies. Results for subject SR2 are presented in Fig. 8. Electrode 3 was stimulated at MCL levels for all stimuli, which included 200 ms bursts of both SAM and unmodulated pulses. Six conditions were studied for each of five carrier rates and for the unmodulated pulse trains. The conditions for the SAM pulse trains included 100 percent modulation at frequencies of 100 through 600 Hz at 100 Hz intervals. The conditions for the unmodulated pulse trains included the corresponding pulse rates of 100 through 600/s. Separate scaling experiments were conducted for each of the carrier rates and for the unmodulated pulse trains. The amplitudes of the stimuli were adjusted prior to each experiment as necessary to eliminate any differences in loudness across conditions. The subject was instructed to assign

a number between 0 and 100 for each stimulus in the experiment according to perceived pitch. Thirty stimuli per condition were presented in random order across conditions for the experiments involving unmodulated pulses and the carrier rates of 2032, 5081 and 10162 pulses/s. Sixty stimuli per condition were presented for the experiments involving the carrier rates of 504 and 1016 pulses/s.

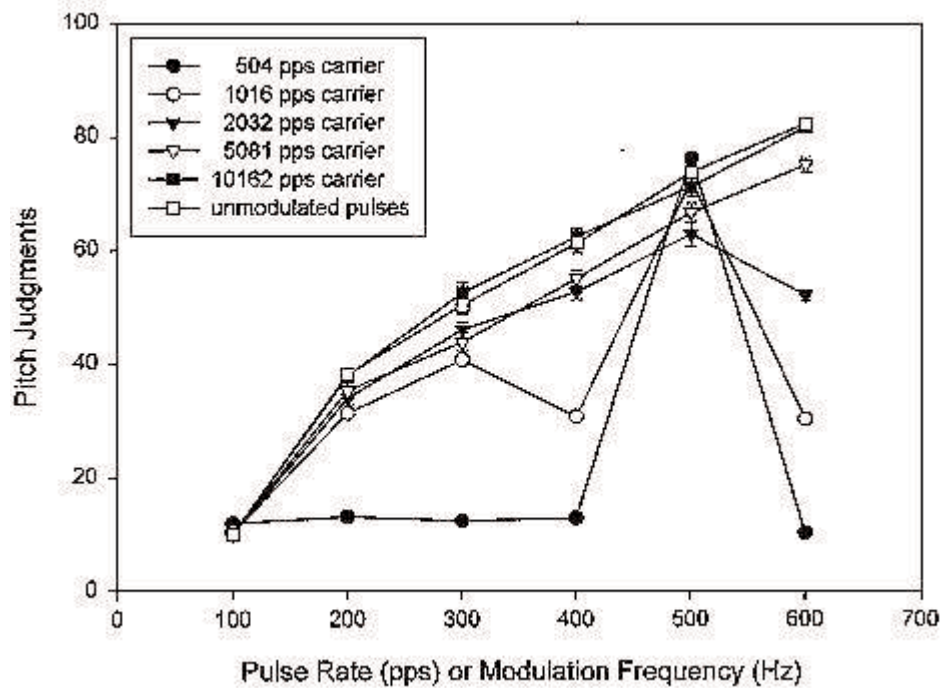


Fig. 8. Scaling of pitch judgments for unmodulated pulse trains as a function of pulse rate, and for SAM pulse trains as a function of modulation frequency. Data are from studies with subject SR2.

Figure 8 shows the means and standard errors of the means of the judgments for each of the six conditions within the six experiments. As expected from the prior EP recordings, carrier rate influenced the range over which pitch increased monotonically with increases in modulation frequency. For the 1016 pulses/s carrier, increases in modulation frequency beyond 300 Hz did not produce monotonic increases in pitch. In fact, judged pitch is not statistically different for the 200, 400 and 600 Hz modulation conditions. This is consistent with a predominance of 5 ms intervals between major peaks in the neural response patterns for these conditions, as shown in the left column of Fig. 7. The judgment for 500 Hz modulation is substantially higher than the judgment for all other conditions. This again is consistent with the pattern of neural responses, which shows peaks separated by 2 ms. The increases in pitch up to the modulation frequency of 300 Hz may correspond to a progressive reduction in the intervals between principal peaks in the neural response patterns as the frequency is increased from 100 to 300 Hz.

Judgments for the 504 pulses/s carrier are highly similar for the modulation frequencies of 100, 200, 300, 400 and 600 Hz. In fact, the judgments for the lower four frequencies are statistically identical. Although this result may seem curious at first sight, plots of the stimuli show that each of these particular combinations of modulation frequency and carrier rate produce a predominance of 10 ms intervals between major peaks.

The judgment for the 500 Hz modulation condition and the 504 pulses/s carrier shows a large increase in pitch compared with the judgments for the other modulation conditions. The close approximation of the modulation frequency to the carrier rate produces peaks in the stimuli at 2 ms intervals.

Pitch increases monotonically with increases in modulation frequency up to 500 Hz for the 2032 pulses/s carrier. Pitch is reduced for the 600 Hz modulation condition, and this judgment does not differ significantly from the judgment for the 400 Hz modulation condition.

For higher carrier rates, and for unmodulated pulses, pitch scales monotonically with increases in modulation frequency or pulse rate, respectively. The range of pitch judgments is greatest for the highest carrier rate and for the unmodulated

pulses.

Discussion

Results like those of Fig. 8 show that increases in carrier rate can increase the range over which increases in modulation frequency produce monotonic increases in pitch. The question now is whether access to a greater range of pitches on single channels will be helpful in multichannel processors. One concern is that temporal channel interactions, due to summation at neural membranes of rapidly presented pulses from different electrodes, may be exacerbated with increased rates of stimulation. Such interactions may limit possible benefits of high rate stimuli. On the other hand, control or reduction of temporal channel interactions might remove this potential limitation. Work is in progress to develop novel stimulus waveforms specifically designed to leave neural membranes at their resting potential after delivery of a subthreshold pulse (e.g., Eddington *et al.*, 1994). In this way, effects of sequential subthreshold pulses (from different electrodes) would not accumulate and cause an unwanted discharge of the neuron. In addition, parallel developments of new electrode designs (e.g., Kuzma, 1996; Kuzma *et al.*, 1996; Seldon *et al.*, 1994) may produce an electrode system with greater spatial specificity of stimulation and reduced interactions among electrodes compared to present designs. Use of such electrodes may reduce or eliminate the concern about temporal channel interactions.

Implementation of high rate processors with multiple channels is technically demanding. Delivery of high rate stimuli within and across electrodes requires high bandwidth current sources capable of generating pulses with durations of 10 μ s/phase or less. Also, for percutaneous connector systems, capacitive coupling among the leads in the cable to the connector can lead to "crosstalk" among channels for high frequency, high rate stimuli. New equipment has been developed in our laboratory (van den Honert *et al.*, 1996) and at the Massachusetts Eye and Ear Infirmary to support evaluations of high rate processors. In our system, for example, each of 24 current sources can generate pulses with phase durations as short as 5 μ s, and capacitive crosstalk among the leads to percutaneous connectors has been reduced to insignificant levels through the use of "driven" shields for each of the leads. Results from the initial studies in our laboratory should be available before the end of 1997.

We note that results of recent studies by Brill *et al.* (in press), Brill and Hochmair (1997), and Kiefer *et al.* (1996; 1997) 1997) have provided some preliminary indications of improved speech reception scores with high rate stimuli. In the near future we should know whether further improvements can be obtained with combinations of numbers of channels and rates of stimulation that exceed the range of the implant systems used in the above studies, which are also the fastest implant systems commercially available at this time. In addition, further improvements may be obtained through different choices of processor parameters, such as an increase in the cutoff frequency of the lowpass filters in the envelope detectors of CIS processors, to make available to patients higher frequencies in modulation waveforms.

References

- Brill S, Gstöttner W, Helms J, von Ilberg C, Baumgartner W, Müller J, Kiefer J. Optimization of channel number and stimulation rate for the fast CIS-strategy in the COMBI 40+. *Am J Otol*, in press.
- Brill S, Hochmair ES. Speech understanding as a function of the number of active channels and stimulation rate in the CIS strategy as implemented in the Combi 40/Combi 40+. Presented at the Vth International Cochlear Implant Conference, New York, NY, May, 1997.
- Busby PA, Tong YC, Clark GM. The perception of temporal modulations by cochlear implant patients. *J Acoust Soc Am* 1993; 94: 124-131.
- Eddington DK, Runbinstein JT, Dynes SBC. Forward masking during intracochlear electrical stimulation: Models, physiology, and psychophysics. *J Acoust Soc Am* 1994; 95: 2904.
- Hartmann R, Topp G, Klinke R. Discharge patterns of cat primary auditory nerve fibers with electrical stimulation of the cochlea. *Hear Res* 1984; 13: 47-62.
- Kiefer J, Pfennigdorff T, Rupprecht V, Huber-Egener J, von Ilberg C. The effect of stimulus rate and channel number on speech understanding with the CIS-strategy in cochlear implant patients. Presented at the Third European

Symposium on Paediatric Cochlear Implantation, Hannover, Germany, June, 1996.

Kiefer J, von Ilberg C, Rupprecht V, Huber-Egener J, Baumgartner W, Gstöttner W, Stephan K. Optimized speech understanding with the CIS-speech coding strategy in cochlear implants: The effect of variations in stimulus rate and numbers of channels. Presented at the Vth International Cochlear Implant Conference, New York, NY, May, 1997.

Kuzma, JA. Cochlear electrode implant assemblies with positioning system therefor. US patent 5545219, Aug. 13, 1996.

Kuzma JA, Seldon HL, Brown GG. Self-curving cochlear electrode array. US patent 5578084, Nov. 26, 1996.

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

Parnas BR. Noise and neuronal populations conspire to encode simple waveforms reliably. *IEEE Trans Biomed Eng* 1996; 43: 313-318.

Rabiner LR, Schafer RW. Digital processing of speech signals. Englewood Cliffs, NJ: Prentice-Hall, 1978.

Seldon HL, Daham MC, Clark GM. Silastic with polyacrylic acid filler: Swelling properties, biocompatibility and potential use into cochlear implants. *Biomaterials* 1994; 15: 1161-1169.

van den Honert C, Zerbi M, Finley C, Wilson B. Speech processors for auditory prostheses. Fourth Quarterly Progress Report, NIH project N01-DC-5-2103. Neural Prosthesis Program, National Institutes of Health, Bethesda, MD, 1996.

Wilson BS, Finley CC, Lawson DT, Zerbi M. Speech processors for auditory prostheses. Eleventh Quarterly Progress Report, NIH project N01-DC-2-2401. Neural Prosthesis Program, National Institutes of Health, Bethesda, MD, 1995.

Wilson BS, Finley CC, Lawson DT, Zerbi M. Temporal representations with cochlear implants. *Am J Otol*, in press.

Wilson BS, Finley CC, Lawson DT, Zerbi M, van den Honert C. High rate coding strategies. Presented at the International Workshop on Cochlear Implants, Vienna, Austria, Oct., 1996.

Wilson BS, Finley CC, Zerbi M, Lawson DT. Speech processors for auditory prostheses. Seventh Quarterly Progress Report, NIH project N01-DC-2-2401. Neural Prosthesis Program, National Institutes of Health, Bethesda, MD, 1994.

III. Plans for the Next Quarter

Our plans for the next quarter include the following:

- Presentation of four invited lectures, including one at the Vth International Cochlear Implant Conference in New York City, May 1-2, two at the 131st annual meeting of the American Otological Society in Scottsdale, AZ, May 10-11, and one at Johns Hopkins University, May 30.
- Ongoing studies with Ineraid subject SR2 (usually one morning per week). Studies anticipated for the next quarter include various psychophysical, electrophysiological and speech reception measures to evaluate further the potential efficacy of high rates of stimulation for cochlear implants.
- Studies with Ineraid subjects SR9, SR15 and SR16, principally to begin evaluation of possible learning effects with chronic use of a portable CIS processor. Each subject will be fit with the variation of the Innsbruck processor developed with our assistance. Additional speech reception, psychophysical, and evoked potential studies are planned for all three subjects. SR15 is scheduled to visit RTI the week beginning on June 2, SR16 is scheduled for the week beginning on June 16, and SR9 is scheduled for the week beginning on July 28.
- Studies with Ineraid subject SR3, including (a) a repeated set of tests to evaluate possible continuing learning effects with a portable CIS processor first applied for this subject in April, 1995, and last evaluated in January, 1996; (b) evaluation of high rate speech processors; (c) recordings of intracochlear EPs to high rate stimuli; (d) recordings of EABRs to 50 pulses/s stimuli for all six electrodes; (e) cross-calibration of new consonant test

materials; and (f) psychophysical measures of pitch scaling of modulation frequencies for sinusoidally amplitude modulated (SAM) pulse trains and of pulse rate for unmodulated pulse trains. The SAM pulse trains for the latter measures will include carrier rates ranging from 504 to 10162 pulses/s. SR3 is scheduled to visit RTI for the two weeks beginning on July 7.

- Continued preparation for studies with recipients of bilateral implants.
- Resumption of studies with subject NU-4, a local patient with standard transcutaneous Nucleus implants on both sides (this subject has made arrangements to work with us for up to two half days per week throughout the summer and fall of 1997).
- Continued development of a new type of compression function for use in CIS processors, designed to reproduce principal features of the noninstantaneous compression found in normal hearing at the synapse between sensory hair cells and adjacent neurons.
- Continued preparation for psychophysical studies to investigate pitch perception with structured random pulse trains and to investigate effects of pulse polarity, number and rate on the threshold of audibility.
- Continued analysis of speech reception and evoked potential data from prior studies.
- Continued preparation of manuscripts for publication.

IV. Acknowledgments

We thank subjects SR2, SR10, NP2 and NP4 for their participation in the studies of this quarter.

Appendix 1. Summary of Reporting Activity for this Quarter

Reporting activity for the last quarter, covering the period of February 1 through April 30, 1997, included the following presentations:

van den Honert C: Microstimulation of auditory nerve for estimating cochlear place of single fibers in a deaf ear. Invited lecture, Department of Otolaryngology, University of Iowa Hospitals and Clinics, Iowa City, IA, April 21, 1997.

Lawson DT: Cochlear implant research. Invited lecture, Duke University Medical Center symposium "Excellence in Otolaryngology – Head & Neck Surgery," Durham, NC, April 25 and 26, 1997.