X
B

Final Report

May 1, 1989 through July 31, 1992

NIH Contract N01-DC-9-2401

Speech Processors for Auditory Prostheses

Prepared by

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

Neuroscience Program Office
Research Triangle Institute
Research Triangle Park, NC 27709

CONTENTS

I. Introduction

The purpose of this project was to design and evaluate speech processors for implantable auditory prostheses. Ideally, such processors extract (or preserve) from speech those parameters that are essential for intelligibility and then appropriately represent these parameters for electrical stimulation of the auditory nerve or central auditory structures.

A principal achievement of this project was discovery of a new way of presenting speech information that produces remarkable gains in speech recognition for implant patients. Design and evaluation of this *continuous interleaved sampling* (CIS) strategy are described in Quarterly Progress Reports (QPRs) 2, 4, 5, 11 and 12. QPR 4 describes comparisons of the CIS and *compressed analog* (CA) strategies in tests with seven Ineraid patients who had high levels of performance with their clinical CA processors, and the subsequent QPRs describe comparisons of those strategies in tests with four additional patients who had poor performance with the clinical device. QPR 2 describes an initial evaluation of CIS processors with one of the "high performance" subjects (Ineraid subject SR2). Every one of these eleven subjects enjoyed immediate, substantial gains in open-set speech recognition when the CIS strategy was substituted for the CA strategy. For example, one of the subjects who started with low levels of clinical performance achieved scores with the CIS processor that, with the CA, would have qualified him for membership in the high performance group.

Results from the first seven subjects have been published in *Nature*.

In addition to the work mentioned above, we have:

- Evaluated in a preliminary way several other promising strategies, including the *peak picker* (PP) strategy of QPR 3 and the hybrid PP/CIS strategy of QPR 10.
- Demonstrated open-set speech recognition with the Auditory Brainstem Implant, using a pulsatile single-channel *continuous sampling* (CS) strategy (see QPRs 6 and 12).
- Completed a series of comparisons between the CA and *interleaved pulses* (IP) strategies (see QPR 1 and Wilson et al., 1991b).
- Evaluated limited implementations of the CIS strategy in studies with patients implanted with the Nucleus and UCSF/Storz devices.
- Completed the prototype for a portable processor using two DSP56001 digital signal processing chips (see QPRs 7 and 9).
- Initiated studies of speech reception in noise with the CA and CIS processors (see QPR 10 and Lawson et al., 1992).
- Conducted preliminary psychophysical studies of frequency discrimination and scaling.
- Conducted preliminary psychophysical studies of perception of the modulation waveform and carrier in sinusoidally modulated pulse trains (see QPR 9).
- Measured dynamic ranges for a variety of pulse durations and rates (see QPR 9).
- Evaluated the new MiniMed speech processor, transmission system, and receiver system in tests with several subjects (see QPR 9 for a description of tests and results with one of the subjects).
- Initiated systematic studies of parameter manipulations in the CS and CIS processors (see QPRs 6, 9, 10, 11 and 12, and see Lawson et al., 1992).

- Evaluated correlations of within-subject speech recognition scores using different processing strategies (see QPR 8 and Wilson et al., 1992).
- Developed hardware, software, and testing procedures to support the above studies.
- Published 10 papers, presented 16 invited lectures at national and international conferences, published 8 abstracts, and presented 5 other papers at national and international conferences.

In this report we present results from several of the studies and activities outlined above. We also list publications and presentations for the project, and briefly indicate directions for future research. Additional information may be found in the QPRs and papers cited above, and in the publications listed in section VI. Several reprints, along with copies of all papers in press, are included as appendices to this report.

II. Comparisons of CA and CIS Processors for Multichannel Cochlear Implants

Recent studies in our laboratory have focused on comparisons of *compressed analog* (CA) and *continuous interleaved sampling* (CIS) processors (Lawson et al., 1992; Wilson et al., 1990b and 1991a). Both use multiple channels of intracochlear electrical stimulation, and both represent waveforms or envelopes of speech input signals. No specific features of the input, such as the fundamental or formant frequencies, are extracted or explicitly represented. CA processors use continuous analog signals as stimuli, whereas CIS processors use nonsimultaneous pulses. The CA approach is used in the widely-applied Ineraid device (Eddington, 1980 and 1983) and in the now-discontinued UCSF/Storz device (with some differences in details of processor implementation, see Merzenich et al., 1984). Wearable devices capable of supporting the CIS approach are just becoming available for use in clinical settings.

We have completed a study of eleven subjects -- seven selected for their high levels of speech recognition with the Ineraid CA processor and four selected for their relatively poor performances with that processor. The "high performance" subjects were representative of the best results, in terms of speech recognition scores, obtained with any commercially-available implant system (Wilson et al., 1991a). The purpose of this section is to provide a summary of results for both sets of subjects.

Processing Strategies

Distinctions between CA and CIS processors are illustrated in Figs. 1 and 2. In CA processors a microphone signal varying over a wide dynamic range is compressed or restricted to the narrow dynamic range of electrically-evoked hearing (Pfingst, 1984; Shannon, 1983) using an automatic gain control. The resulting signal then is filtered into four contiguous frequency bands for presentation to each of four electrodes. As shown in Fig. 1, information about speech sounds is contained in the relative stimulus amplitudes among the four electrode channels and in the temporal details of the waveforms for each channel.

A concern associated with this method of presenting information is that substantial parts of it may not be perceived by implant patients (Wilson et al., 1990a). For example, most patients cannot perceive frequency changes in stimulus waveforms above about 300 Hz (see, e.g., Shannon, 1992). Thus, many of the temporal details present in CA stimuli are not likely to be accessible to the typical user.

In addition, the simultaneous presentation of stimuli may produce significant interactions among channels through vector summation of the electric fields from each electrode (e.g., White et al., 1984). The resulting degradation of channel independence would be expected to reduce the salience of channel-related cues. That is, the neural response to stimuli from one electrode may be significantly distorted, or even counteracted, by coincident stimuli from other electrodes.

The CIS approach addresses the problem of such channel interactions through the use of interleaved nonsimultaneous stimuli (Fig. 2). Trains of balanced biphasic pulses are delivered to each electrode with temporal offsets that eliminate any overlap across channels. The amplitudes of the pulses are

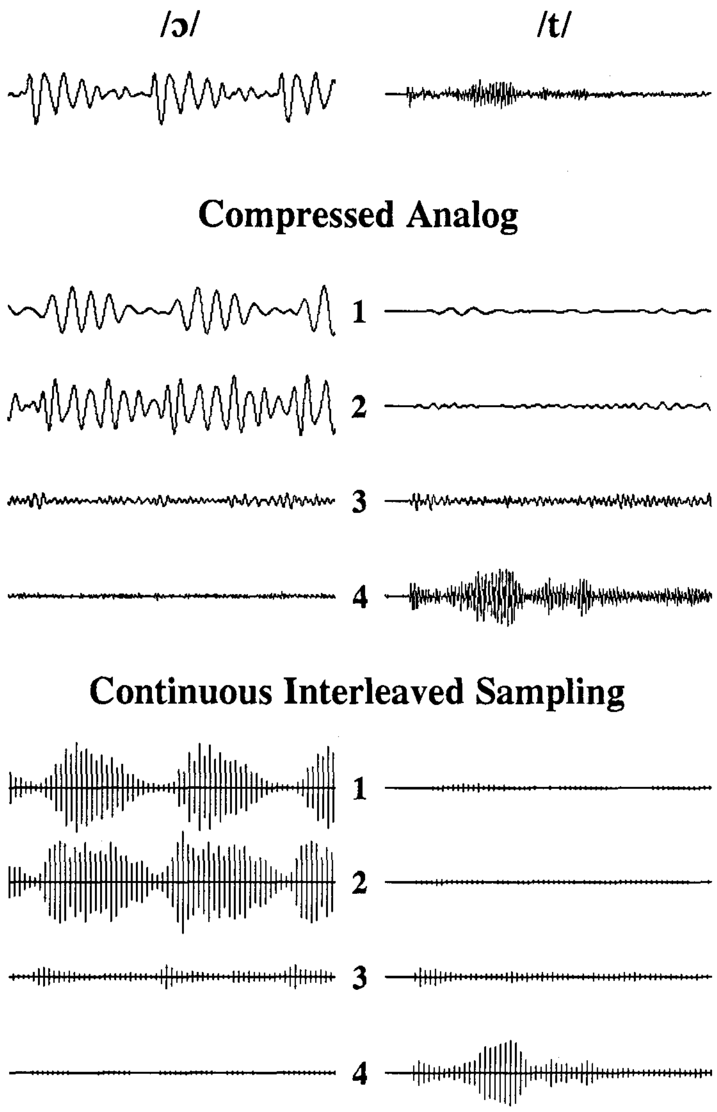


Figure 1. Waveforms produced by simplified implementations of CA and CIS strategies. The top panel shows preemphasized (6 dB/octave attenuation below 1.2 kHz) speech inputs. Inputs corresponding to a voiced speech sound ("aw") and an unvoiced speech sound ("t") are shown in the left and right columns, respectively. The duration of each trace is 25.4 ms. The remaining panels show stimulus waveforms for CA and CIS processors. The waveforms are numbered by channel, with channel 1 delivering its output to the apical-most electrode. To facilitate comparisons between strategies, only four channels of CIS stimulation are illustrated here. In general, five or six channels have been used for that strategy. The pulse amplitudes reflect the envelope of the bandpass output for each channel. In actual implementations the range of pulse amplitudes is compressed using a logarithmic or power-law transformation of the envelope signal.

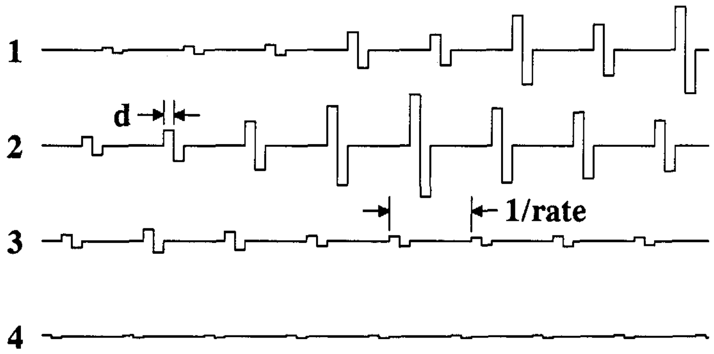


Figure 2. Expanded display of CIS waveforms. Pulse duration per phase ("d") and the period between pulses on each channel ("1/rate") are indicated. The sequence of stimulated channels is 4-3-2-1. The total duration of each trace is 3.3 ms.

derived from the envelopes of bandpass filter outputs. In contrast to the four-channel clinical CA processors, five or six bandpass filters (and channels of stimulation) generally have been used in CIS systems to take advantage of additional implanted electrodes and reduced interactions among channels. The envelopes of the bandpass outputs are formed by rectification and lowpass filtering. Finally, the amplitude of each stimulus pulse is determined by a logarithmic or power-law transformation of the corresponding channel's envelope signal at that time. This transformation compresses each signal into the dynamic range appropriate for its channel.

A key feature of the CIS approach is its relatively high rate of stimulation on each channel. Other pulsatile strategies present sequences of interleaved pulses across electrodes at a rate equal to the estimated fundamental frequency during voiced speech and at a jittered or fixed (often higher) rate during unvoiced speech (Clark, 1987; Wilson, 1992; Wilson et al., 1991b). Rates of stimulation on any one channel rarely have exceeded 300 pulses per second (pps). In contrast, CIS processors generally use brief pulses and minimal delays, so that rapid variations in speech can be tracked by pulse amplitude variations. The rate of stimulation on each channel usually exceeds 500 pps and is constant during both voiced and unvoiced intervals. A constant high rate allows relatively high cutoff frequencies for the lowpass filters in the envelope detectors. With a stimulus rate of 800 pps, for instance, lowpass cutoffs can approach (but not exceed) 400 Hz without introducing aliasing errors in the sampling of the envelope signals at the time of each pulse (see Rabiner and Shafer, 1978, for a complete discussion of aliasing and its consequences).

Methods

Each subject has been studied for a one-week period during which (a) basic psychophysical measures were obtained on thresholds and dynamic ranges for pulsatile stimuli, (b) a variety of CIS processors (with different choices of processor parameters) were evaluated with preliminary tests of consonant identification, and (c) performance with the best of the CIS processors and the clinical CA processor was documented with a broad spectrum of speech tests. Experience with the clinical processor exceeded one year of daily use for all subjects. In contrast, experience with the CIS processors was limited to no more than several hours before formal testing. All comparisons within this eleven-subject study are on the basis of a single week of CIS optimization. In subsequent visits by some of the same subjects a potential for significant further optimization has been demonstrated.

Tests. The comparison tests included open-set recognition of 50 one-syllable words from Northwestern University Auditory Test 6 (NU-6), 25 two-syllable words (spondees), 100 key words in the Central Institute for the Deaf (CID) sentences of everyday speech, and the final word in each of 50 sentences from the Speech Perception in Noise (SPIN) test (presented in our studies without noise). All tests were conducted with hearing alone, using single presentations of recorded material, and without feedback as to correct or incorrect responses.

Processor parameters. Each subject's own clinical device was used for the tests with the CA processor. As mentioned above, selection of parameters for the CIS processor was guided by preliminary tests of consonant identification. The standard four channels of stimulation were used for the clinical CA processors (Eddington, 1980 and 1983), whereas five or six channels were used for the CIS processors. Additional parameters of the CIS processors are presented in Table 1. As indicated there, all CIS processors for the "high performance" subjects, SR2-8, had pulse durations of 102 μs/phase or less, zero delay between the sequential pulses on different channels, pulse rates of 817 pps or higher on each channel, and a cutoff frequency for the lowpass filters of 400 Hz or higher. The best processor for subject SR1 also fit this description, except that a delay of 172 μs was interposed between sequential pulses. The best processors for subjects SR9-11 used long-duration pulses (167 μs/phase), paired with a relatively low rate of stimulation on each channel (500 pps) and a relatively low cutoff frequency for the lowpass filters (200 Hz).

Evaluation of practice and learning effects. Because the tests with the CA processor preceded those with the selected CIS processor for each subject, we were concerned that practice or learning effects might favor the latter in comparisons of the two strategies. To evaluate this possibility, the CID and NU-6 tests were repeated with the CIS processor for five of the "high performance" subjects (subjects SR3, SR4 and SR6-8), using a different recorded speaker and new lists of words and sentences. Practice or learning effects would be demonstrated by significant differences in the test/retest scores. However, no such differences were found ($p > 0.6$ for paired t comparisons of the CID scores; $p > 0.2$ for the NU-6 scores), and the scores from the first and second tests were averaged for all subsequent analyses.

Table 1. Parameters of CIS processors. The parameters include number of channels, pulse duration, the rate of stimulation on each channel (Rate), and the cutoff frequency of the lowpass integrating filters for envelope detection (Integrating Filter Cutoff). The subjects are listed in the chronological order of their participation in the present studies. SR2 through SR8 are the "high performance" subjects while SR1 and SR9-11 belong to the "low performance" group.

Subject	Channels	Pulse Duration (μs/phase)	Rate (pps)	Integrating Filter Cutoff (Hz)
SR2	6	55	1515	800
SR3	6	31	2688	400
SR4	6	63	1323	400
SR5	6	31	2688	800
SR6	6	102	817	400
SR7	5	34	2941	400
SR8	6	100	833	400
SR1	5	34	833	400
SR10	6	167	500	200
SR9	5	167	500	200
SR11	6	167	500	200

Results

The results from one-week studies of each of the eleven subjects are presented in Table 2 and Fig. 3. CA and CIS scores for each of the "high performance" subjects are connected by the light lines near the top of each panel in Fig. 3, and scores for the four "low performance" subjects are connected by the dark lines closer to the bottom of each panel. We note that low-performance subject SR1 had participated in an earlier study not involving CIS processors (Wilson et al., 1991b). Results from his first week of testing with CIS processors are presented here. This is also true of high-performance subject SR2, who has returned to the laboratory for many additional studies with various implementations of CIS processors (see, e.g., Lawson et al., 1992). In those subsequent tests SR2 has achieved even higher scores using a variety of six-channel CIS processors, with NU-6 percentages ranging from the high 80s to the low 90s.

As is evident from the figure, scores for all eleven subjects are improved with the use of a CIS processor. The average scores across subjects increased from 57 to 80% correct on the spondee test ($p < 0.002$), from 62 to 84% correct on the CID test ($p < 0.005$), from 34 to 65% correct on the SPIN test ($p < 0.001$), and from 30 to 47% correct on the NU-6 test ($p < 0.0005$). Note that the range of difficulty among our four tests provides sensitivity to performance differences across the rather wide range of absolute performance represented in this eleven-subject study.

Table 2. Individual results from the open-set tests.

	Spondee		CID		SPIN		NU-6	
Subject	CA	CIS	CA	CIS	CA	CIS	CA	CIS
SR2	92	96	100	100	78	96	56	80
SR3	52	96	66	98	14	92	34	58
SR4	68	76	93	95	28	70	34	40
SR5	100	100	97	100	94	100	70	80
SR6	72	92	73	99	36	74	30	49
SR7	80	100	99	100	66	98	38	71
SR8	68	100	80	100	36	94	38	66
SR1	40	60	25	70	2	30	6	32
SR10	0	56	1	55	0	26	0	14
SR9	8	34	9	34	2	2	2	4
SR11	46	66	40	71	12	30	18	22

Perhaps the most encouraging of these results are the improvements for the four low-performance subjects. SR1, for instance, achieved scores with the CIS processor that would have qualified him for membership in the high performance group (with the clinical CA processor). Similarly, SR10 achieved relatively high scores with the CIS processor. The score on the spondee test increased from 0 to 56% correct, on the CID test from 1 to 55% correct, on the SPIN test from 0 to 26% correct, and on the NU-6 test from 0 to 14% correct. These increases were obtained with no more than several hours of aggregated experience with CIS processors, compared to more than a year of daily experience with the clinical CA processor.

Note that while these gains for SR10 are large, they are not atypical of results for the other subjects. His improvements follow the pattern of the other subjects, i.e., generally large gains in the scores of tests that are not limited by ceiling effects. The distinctive aspect of SR10's results is that he enjoys such gains even though he started at or near zero on all four tests. Thus, the relative improvements for SR10 are larger than those for any other subject in the series.

Discussion

The findings presented above demonstrate that use of CIS processors can produce large and immediate gains in speech recognition for a wide range of implant patients. Indeed, the sensitivity of some of the administered tests has been limited by ceiling (or saturation) effects: five of the seven "high performance" subjects scored 96% or higher for the spondee test using CIS processors; all seven scored 95% or higher for the CID test; and five scored 92% or higher for the SPIN test. Scores for the NU-6

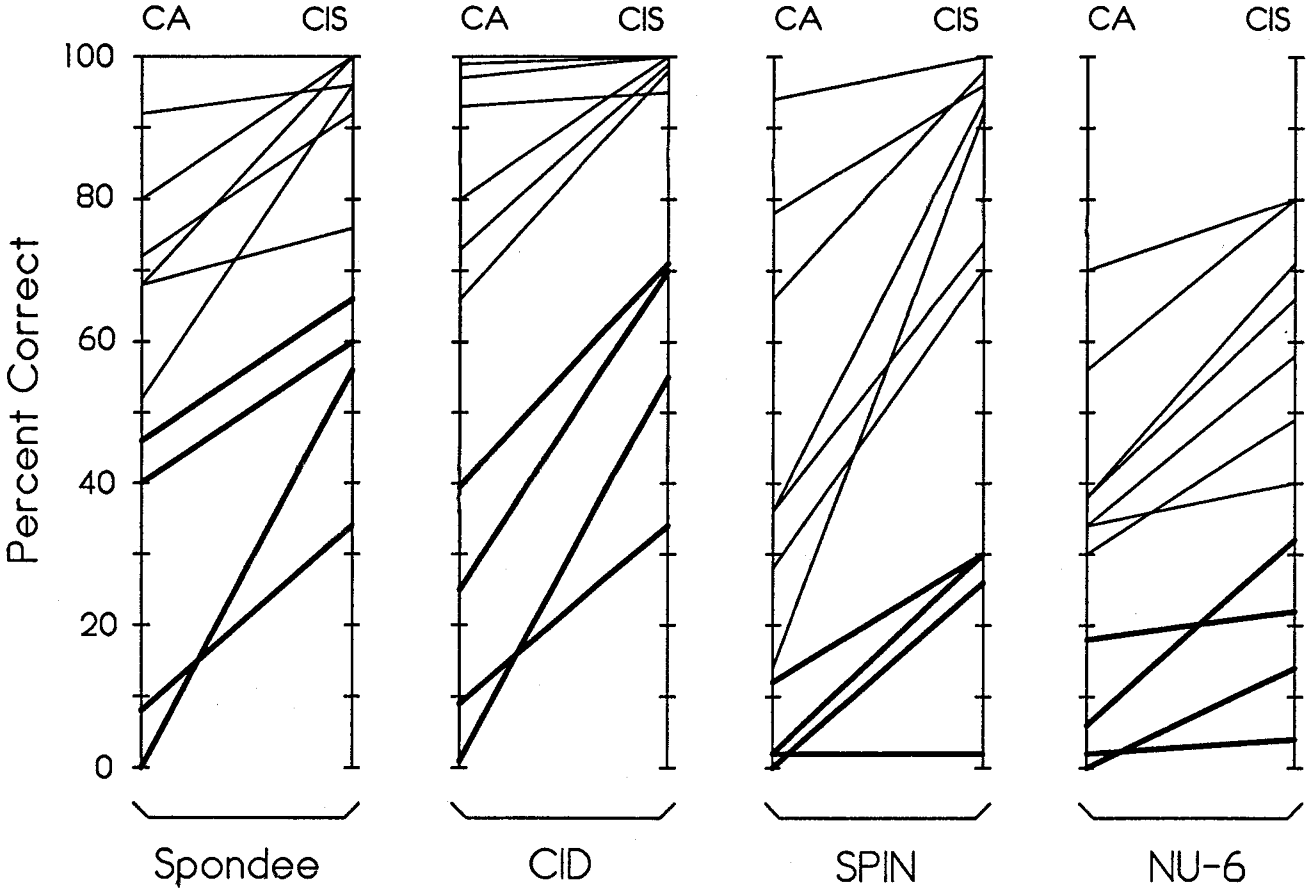


Figure 3. Speech recognition scores for CA and CIS processors. A line connects the CA and CIS scores for each subject. Light lines correspond to the seven subjects selected for their excellent performance with the clinical CA processor, while the heavier lines correspond to the four subjects selected for relatively poor performance.

test, while not approaching the ceiling, still were quite high. The 80% score achieved by two of the subjects corresponds to the middle of the range of scores obtained by people with mild-to-moderate hearing losses when taking the same test (Bess and Townsend, 1977; Dubno and Dirks, 1982).

The improvements are even more striking when one considers the large disparity in experience with the two processors. At the time of our tests each subject had 1 to 5 years of daily experience with the CA processor, but only several hours over a few days with CIS. In previous studies involving within-subjects comparisons, such differences in experience have strongly favored the processor with the greatest duration of use (Dowell et al., 1987; Dowell et al., 1990; Tyler et al., 1986).

Factors contributing to the performance of CIS processors might include (a) reduction in channel interactions through the use of nonsimultaneous stimuli, (b) use of five or six channels instead of four, (c) representation of rapid envelope variations through the use relatively high pulse rates, (d) preservation of amplitude cues with channel-by-channel compression, and (e) the shape of the compression function.

An interesting aspect of the studies with low-performance subjects is that the best CIS processors seem to involve parameters distinct from those of the best processors for subjects in the high-performance group. The best processor for SR1 used short-duration pulses (34 μs/phase) presented at a relatively low rate (833 pps), and the best processors for SR9-11 used long-duration pulses (167 μs/phase) presented at an even lower rate (500 pps). The subjects in the high-performance group, however, often obtained their best scores with processors tending to minimize pulse widths and maximize pulse rates (e.g., 31 μs/phase pulses presented at 2688 pps).

The use of such shorter pulses and higher rates allows representation of higher frequencies in the modulation waveform for each channel, i.e., the cutoff frequency of the lowpass filter in the envelope detectors for each channel may be raised to 1/2 the pulse rate without introducing aliasing effects. In addition, the dynamic range (DR) of electrical stimulation -- from threshold to most comfortable loudness -- typically is a strong function of pulse rate and a weaker function of pulse duration (Shannon, 1992; Wilson et al., 1991c). Large increases in DR generally are found with increases in pulse rates from about 400 pps to 2500 pps. Smaller increases often (but not always) are observed with increases in pulse duration (at a fixed rate of stimulation) from roughly 50 μs/phase to higher values (e.g., out to 200 μs/phase for practical CIS designs).

For some patients, though, these advantages may be outweighed by other factors. For several subjects in our Ineraid series, for instance, we have observed that the salience of channel ranking can decline with decreases in pulse widths below 100 μs/phase. A favorable tradeoff for such subjects might involve the use of long-duration pulses (e.g., 100 μs/phase or greater) to preserve channel cues, while foregoing any additional DR obtainable with shorter pulses and higher rates of stimulation.

Another possible advantage of relatively low rates of stimulation is further reduction of channel interactions. Providing time between pulses on sequential channels can reduce the "temporal integration" component of channel interactions (a component produced by the accumulation of charge at neural membranes from sequential stimuli, see, e.g., White et al., 1984). Thus, use of time delays between short-duration pulses in the stimulation sequence across electrodes may reduce interactions. Alternatively, use of long-duration pulses with no time delay also might reduce temporal interactions in that a relatively long period still is realized between the excitatory phases of successive pulses.

Collectively the present results indicate that (a) the performance of at least some patients with poor clinical outcomes can be improved substantially with the use of a CIS processor, (b) use of long-duration pulses produced large gains in speech test scores for three such subjects, (c) use of short-duration pulses presented at a relatively low rate produced similar improvements in another such subject, and (d) the optimal tradeoffs among pulse duration, pulse rate, interval between sequential pulses, and cutoff frequency of the lowpass filters appear to vary from patient to patient.

III. Additional Aspects of CIS Performance

In recent studies we have begun systematic investigation of various aspects of CIS performance. These studies include evaluations of CIS performance (a) across numbers of channels, (b) across other manipulations in CIS parameters, and (c) under conditions of noise interference. The purpose of this section is to present preliminary results from studies (a) and (c).

Effects of Manipulations in Channel Number

Both studies were conducted with subject SR2. Because this subject's high scores had compromised the sensitivity of our consonant and other tests in evaluations of CIS processors with a variety of new parameter sets (i.e., he obtained perfect or nearly perfect scores on the consonant, spondee, CID and SPIN tests, and he obtained scores in the high 80s or low 90s on the NU-6 test), we decided to increase the difficulty of our consonant test by increasing the number of consonants from 16 to 24 (the set of 24 includes /b, d, f, g, dʒ, h, j, k, l, m, n, ŋ, p, r, s, ʃ, t, tʃ, ð, θ, v, w, z, ʒ/). Since this change, however, SR2 has achieved scores of 99% correct with the male speaker using five different implementations of CIS processors. Furthermore, scores for the female speaker have been only somewhat lower (as high as 95% correct).

While the sensitivity of even the 24 consonant test now is inadequate to distinguish among the better implementations of CIS processors for subject SR2, the sensitivity of that test is well suited to studies exploring decrements in performance with reduced numbers of channels or with increasing amounts of noise interference.

In the 24 consonant test, multiple exemplars of each consonant were presented in an /a/-consonant-/a/ context from laser videodisc recordings of male and female speakers (Tyler et al., 1987). A single block of trials consisted of five randomized presentations of each consonant by a single speaker. The tests were conducted with hearing alone and without feedback as to correct or incorrect responses.

Results from the first study are presented in Fig. 4. The top panel shows percent-correct scores for CIS processors with 6, 5, 4, 3, and 2 channels, and for an analogous processor with 1 channel (referred to as a CS processor, since interleaving is not applicable to a single channel). The bottom panels show information transmission scores for various articulatory and acoustic features of consonants (Miller and Nicely, 1955). The features include voicing (Voi), nasality (Nsl), frication (Fric), duration (Dur), place of articulation (Plc), and envelope cues (Env).

Each n-channel processor used the n apical-most electrodes and filtered the same total frequency range into n bands of equal width on a logarithmic scale. For example the three channel processor used apical electrodes 1, 2 and 3. All processors used 33 μs/phase pulses, presented at the rate of 2525 pps on each channel (delays were interposed between sequential pulses for processors with fewer than six channels to maintain this constant rate). In addition, each processor used 6th order bandpass filters, fullwave rectifiers, and 400 Hz lowpass filters (1st order). For consistency, a fixed base-to-apex update order was used for all processors. For example, the three channel processor stimulated its electrodes in the sequence 3-2-1.

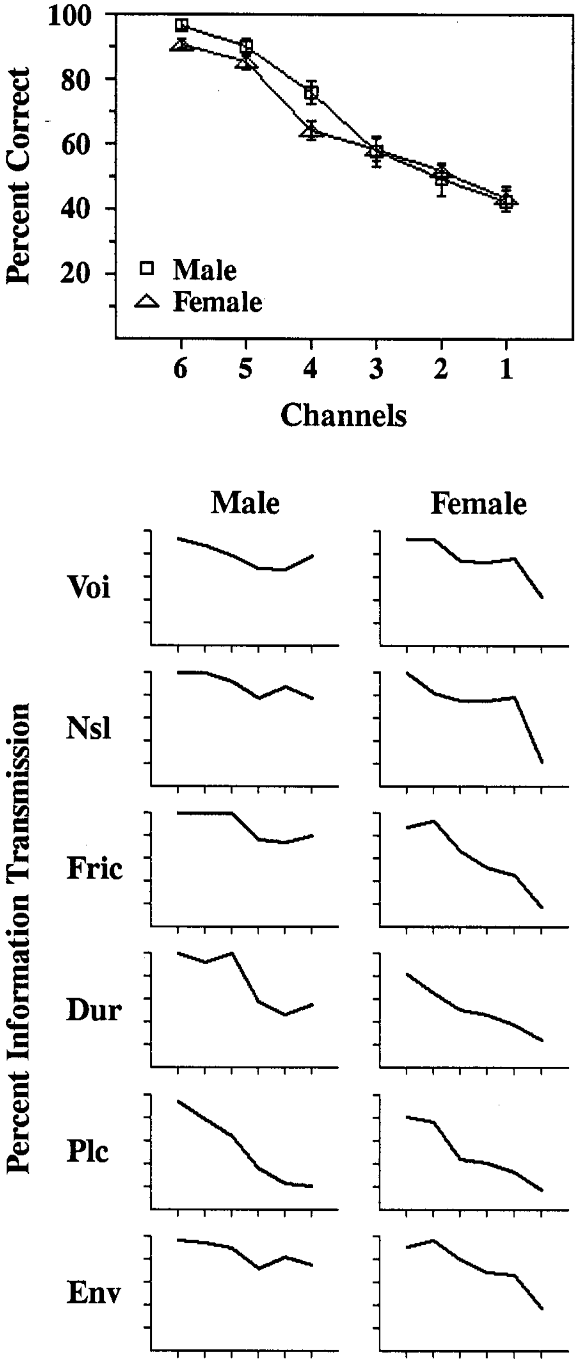


Figure 4. Percent correct and feature transmission scores for processors using different numbers of channels. Five presentations of each of 24 consonants by the male speaker, and five presentations of each consonant by the female speaker, were used in the tests with each processor. The presentations were arranged in block randomized order, providing a percent correct score after each set of randomized presentations of all 24 consonants. The square symbols in the top panel show averages of these scores (from 5 randomized sets) for the male speaker, and the triangles show the averages for the female speaker. Standard errors of the mean are indicated with the vertical bars. The remaining panels show feature transmission scores for the same experimental conditions. Full scale corresponds to 100% information transfer.

We note that none of the processors in this series was optimized for the individual subject. The six-channel version, for instance, was inferior to other six-channel processors using a "staggered" order of channel updates (6-3-5-2-4-1; see Lawson et al., 1992). Also, processors using fewer than six channels probably would have benefited from use of specific electrodes other than the most apical n (e.g., use of more widely spaced electrodes may have produced a better result). The purpose of this particular study was to evaluate effects of changes in the number of channels, while maintaining a consistency in all other CIS parameters.

The results show a strong effect of channel number on consonant identification. Overall percent correct scores decline monotonically, for both the male and female speakers, with reductions in the number of channels. Also, transmission of place information declines precipitously for the male speaker as the number of channels is reduced from 6 to 3, and drops precipitously for the female speaker as the number of channels is reduced from 5 to 4. In all cases the transmission of place information declines monotonically as the number of channels is reduced. In contrast, transmission of envelope information is relatively well maintained when the number of channels is reduced, as is the transmission of voicing, frication and nasality information for the male speaker (indeed, the transmission of voicing information remains high even for a single channel). Results for the female speaker are somewhat different in that the transmission of voicing and nasality information drops sharply when the number of channels is reduced from 2 to 1, and the transmission of frication information drops rapidly over the range of channel reductions from 5 to 1.

A consistent finding in the data is the dependence of place transmission on the number of stimulation channels. In addition, results from the female speaker suggest that transmission of frication information may depend on number of channels, at least up to 5 channels, and at least for certain speakers. Further increases in channel number may improve the transmission of place information and other important cues for the correct identification of consonants. As indicated elsewhere (e.g., Tye-Murray and Tyler, 1989; Dorman et al., 1990; Wilson et al., 1990b), such identification is highly correlated with open-set recognition of words, sentences and running speech.

Effects of Interfering Noise

Results from the second study, to evaluate effects of noise interference, are presented in Fig. 5. Here we show performances of CIS and CA strategies in noise without any special provisions for noise reduction. A six-channel CIS implementation was used with the following parameters: 33 μs/phase pulses, 2525 pps rate of stimulation on each channel, staggered update order, 12th order bandpass filters, fullwave rectifiers, and 400 Hz lowpass filters (1st order).

Consonant identification first was measured under quiet conditions, and then progressively greater amounts of multitalker speech babble were added to the primary speech signal. Signal-to-noise ratios (SNRs) included 15, 10, 5 and 0 dB, with 0 dB corresponding to the babble signal amplitude exceeding the maximum consonant waveform amplitude briefly about once per second on average.

While the presence of noise clearly degrades the performance of both processors, relatively high percent correct scores are maintained down to a SNR of 5 dB. The scores for the CIS processor are

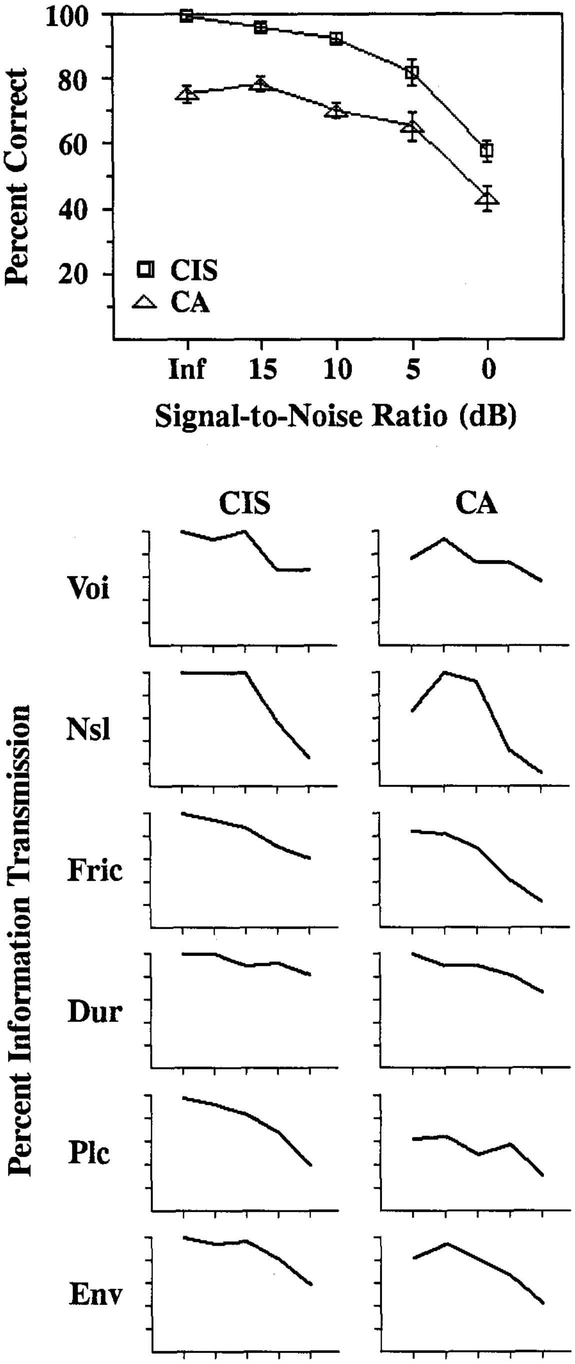


Figure 5. Percent correct and feature transmission scores for CIS and CA processors as a function of signal-to-noise ratio (SNR). The SNR of "Inf" refers to presentation of the signal without any accompanying noise. Five presentations of each of 24 consonants by the male speaker were used in the consonant identification tests for each processor at each SNR. The square symbols in the top panel show average percent correct scores for the CIS processor, and the triangles show the averages for the CA processor. Standard errors of the mean are indicated with the vertical bars. The remaining panels show feature transmission scores for the same experimental conditions. Full scale corresponds to 100% information transfer.

higher than those for the CA processor at all SNRs. This is especially encouraging inasmuch as the CA processor in the Ineraid device has been identified as the most resistant to the deleterious effects of noise among several tested implant systems (Gantz et al., 1987; Tyler and Tye-Murray, 1991).

One possible factor underlying the high levels of CIS performance in the presence of interfering speech babble is a good representation of envelope cues. In particular, covariation in envelope information across channels may help maintain high levels of speech recognition in noise (e.g., Hall et al., 1984; Moore, 1992). Such across-channel information may allow a listener to follow the correlated cues of the primary speech signal, while rejecting the uncorrelated variations produced by the noise.

Another factor that may contribute to the performances found for both the CA and CIS strategies is the fact that neither relies on feature extraction. The accuracy of such extraction can be severely degraded by even modest amounts of noise, as demonstrated in many studies with conventional speech analysis systems (e.g., Rabiner and Shafer, 1978) and in studies with cochlear implant devices (e.g., Gantz et al., 1987).

A key lesson in the present results is that the choice of a basic processing strategy can have large effects on performance in noise.

IV. Evaluation of Other Promising Strategies

While very high levels of speech recognition have been obtained with the CIS strategy, other strategies may well be better, at least for certain classes of patients. One possibility is the *peak picker* (PP) strategy first described in QPR 3 for this project. In studies with one of our Ineraid patients, this strategy produced transmission scores for several consonant features that were higher than the scores obtained with the CIS strategy. Overall transmission of consonant information was approximately the same for the PP and CIS strategies. Transmission of vowel features to this patient by the PP strategy was perfect for our eight vowel test (compared with high, but not perfect, scores for the CIS strategy). The PP strategy obviously has promise and should be investigated further with additional tests and subjects.

A possible advantage of the PP strategy is that it uses generally lower rates of stimulation than the CIS strategy. This may allow useful implementations of the PP strategy for patients implanted with the Nucleus device, whose transcutaneous transmission system (TTS) does not permit the rapid sequencing of pulses typically required to optimize CIS strategy processors.

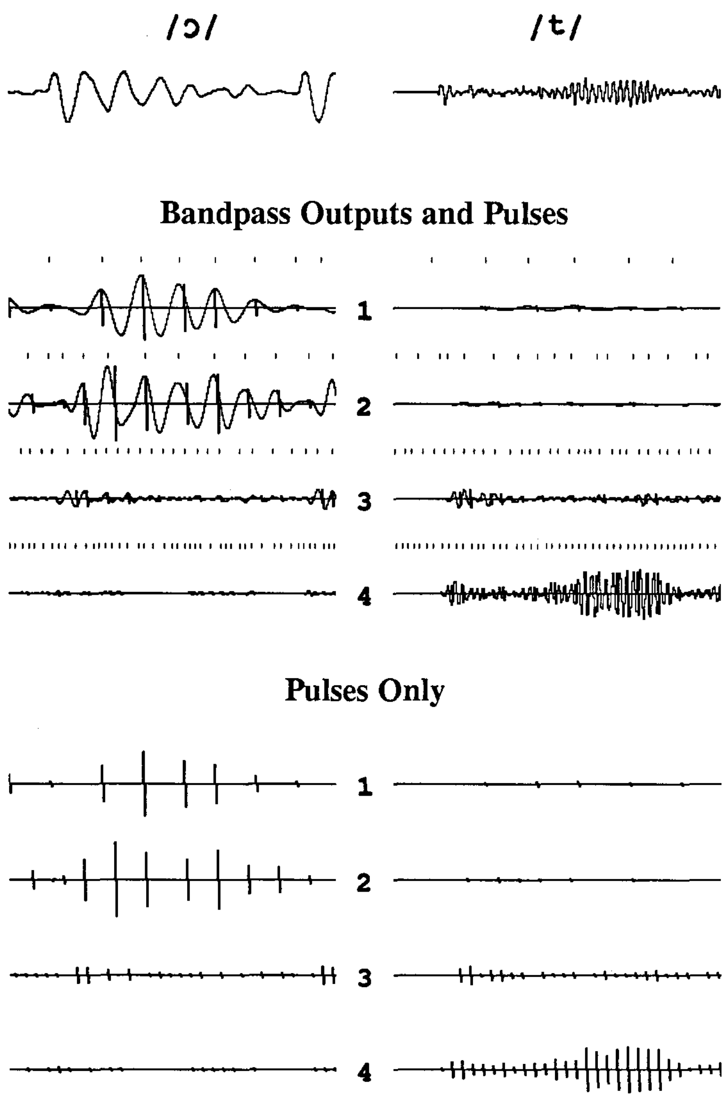
In addition to the PP strategy, we have conducted preliminary studies to evaluate a hybrid PP/CIS strategy. In this strategy PP stimuli are delivered to the apical-most electrodes (usually the two most apical electrodes in an array of six), and CIS stimuli are delivered to the remaining electrodes. This hybrid strategy attempts to combine attributes of the PP and CIS approaches.

Peak Picker (PP) Processor

The design of the PP processor is illustrated in Fig. 6. In this processor the position of a peak in either the bandpass or envelope detector output is signaled by the presentation of a pulse. Also, as in the IP and CIS processors, the exact timing of the pulses is adjusted to avoid any temporal overlap of stimuli across channels.

In Fig. 6 the middle panel shows bandpass outputs for each of four channels along with the stimulus pulses derived from those outputs. In addition, the positions of the peaks in the bandpass outputs are marked by short vertical lines above each trace. The lower panel shows the stimulus pulses only.

In this particular implementation of the PP strategy, each channel is addressed in a fixed sequence and at a fixed rate. A pulse is delivered to a channel if a peak has been detected in its bandpass output since it last was addressed. The amplitude of the pulse is determined with the same logarithmic transformation used in the IP and some CIS processors (i.e., the actual pulse amplitudes would be computed using a logarithmic transformation of the amplitudes shown in the figure). A fixed time is reserved for each channel in the stimulation sequence whether or not a pulse is delivered. As indicated in Fig. 6, this variation of the processor produces clusters of pulses at the F0 rate and individual pulses at the F1 rate for voiced speech sounds (left panels). Because the pulses must be presented nonsimultaneously, though, higher frequencies in the bandpass outputs are not followed with pulses at those frequencies. Notice, for instance, that many peaks are missed in channels 3 and 4, and that large offsets between the positions of peaks and subsequent pulses are seen in the waveforms of channel 2.

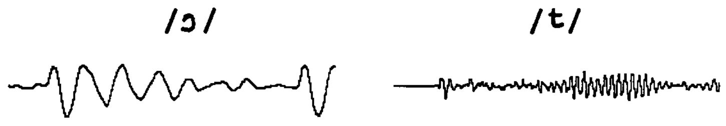


Bandpass Outputs and Pulses

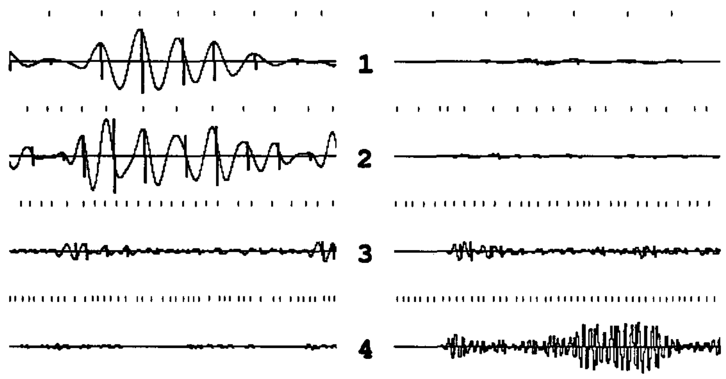
Pulses Only

Figure 6. Waveforms of the "peak picker" (PP) processing strategy. Equalized speech inputs are shown at the top and processor waveforms are shown in the middle and bottom panels. The middle panel shows the bandpass outputs and stimulus pulses for each of four channels. The location of peaks in the bandpass outputs are marked with short vertical lines above each trace. The bottom panel shows the stimulus pulses only. The duration of each trace is 12.25 ms.

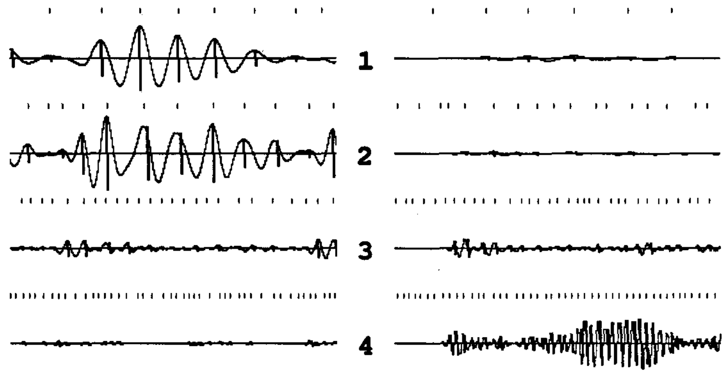
Alternative implementations of the PP processor are illustrated in Fig. 7. The uppermost panel beneath the input signals shows the waveforms of the implementation just described ("Bandpass Outputs, Time for Each Channel"); the next panel down shows an implementation in which the time allocated for each channel is *not* used if a pulse is not to be delivered ("Bandpass Outputs, Channels Skipped"); and the bottom panel shows an implementation in which the outputs of the envelope detectors are used instead



Bandpass Outputs, Time for Each Channel



Bandpass Outputs, Channels Skipped



Envelope Detector Outputs, Channels Skipped

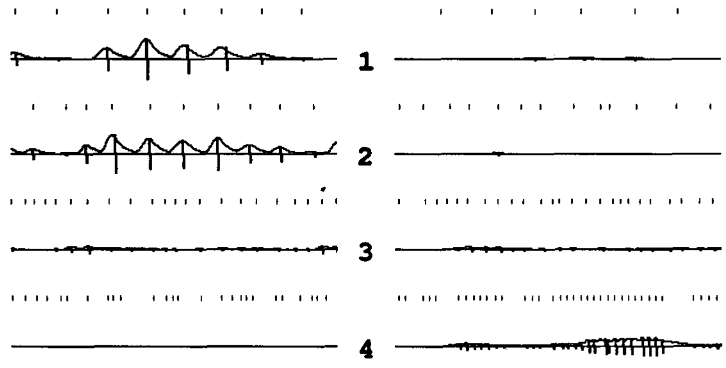


Figure 7. Various implementations of PP processors. See text for details.

of the bandpass outputs ("Envelope Detector Outputs, Channels Skipped").

A summary of waveforms for various types of pulsatile processor is presented in Fig. 8. All three types of processor use nonsimultaneous stimuli. Among these, the CIS processor delivers the greatest number of pulses per unit time, and the IP processor the least. The PP processor provides an intermediate level of temporal detail, with a representation of F1 in the apical channel(s). In addition, the PP processor presents different rates of stimulation on each channel, which might increase the salience of channel-related cues for some patients (i.e., channel cues might be represented both by place of stimulation and by rate of stimulation).

One implementation of a PP processor was evaluated in preliminary tests with Ineraid subject SR2. The design illustrated in Fig. 6 was used, with time taken for each channel whether or not a pulse is delivered. The PP processor used six channels, with the staggered update order found to be best in (contemporaneous) evaluations of CIS processors (6-3-5-2-4-1).

The tests included identification of 16 consonants and 8 vowels, for male and female speakers (see QPR 3 for a complete description of the tests and related procedures).

Results for a contemporaneous implementation of the CIS strategy (processor RTSS8) and the PP strategy (processor RTPP1) are presented in Fig. 9. As noted above, both processors used a staggered update order. In addition, both used pulse durations of 55 μs/phase, no imposed time delay between sequential pulses, and a 600 Hz corner frequency for the input equalization filter (present versions of CIS processors generally use shorter pulses and a 1200 Hz corner frequency for the equalization filter). Finally, the envelope detectors in the CIS processor used halfwave rectifiers and 800 Hz lowpass filters (present CIS processors generally use lower cutoff frequencies for the lowpass filter).

Comparison of speech test scores for the two processors shows a similarity in feature transmission for consonants. Overall transmission is roughly 90% for both strategies. The PP strategy produces somewhat higher scores for the temporal features of voicing, duration, and envelope cues, and the CIS strategy produces somewhat higher scores for the features of nasality and place of articulation. Scores for frication are quite similar for the two processors.

In contrast to the overall picture for consonants, transmission of vowel information appears to be better with the PP processor. In fact, the 8 vowels of our vowel identification test were perfectly identified for both the male and female speakers when the PP processor was used.

The large gain in the transmission of vowel feature information found for F1 is consistent with the explicit representation of F1 in the apical channels with the PP processor.

With the exception of this one feature, and possibly nasality for the consonant test, no obvious differences are found in the results for the two processors. We note, however, that many of the scores approach or encounter the 100% ceiling for these particular tests. More difficult tests will be needed to detect additional differences between the processors, if indeed such differences exist.

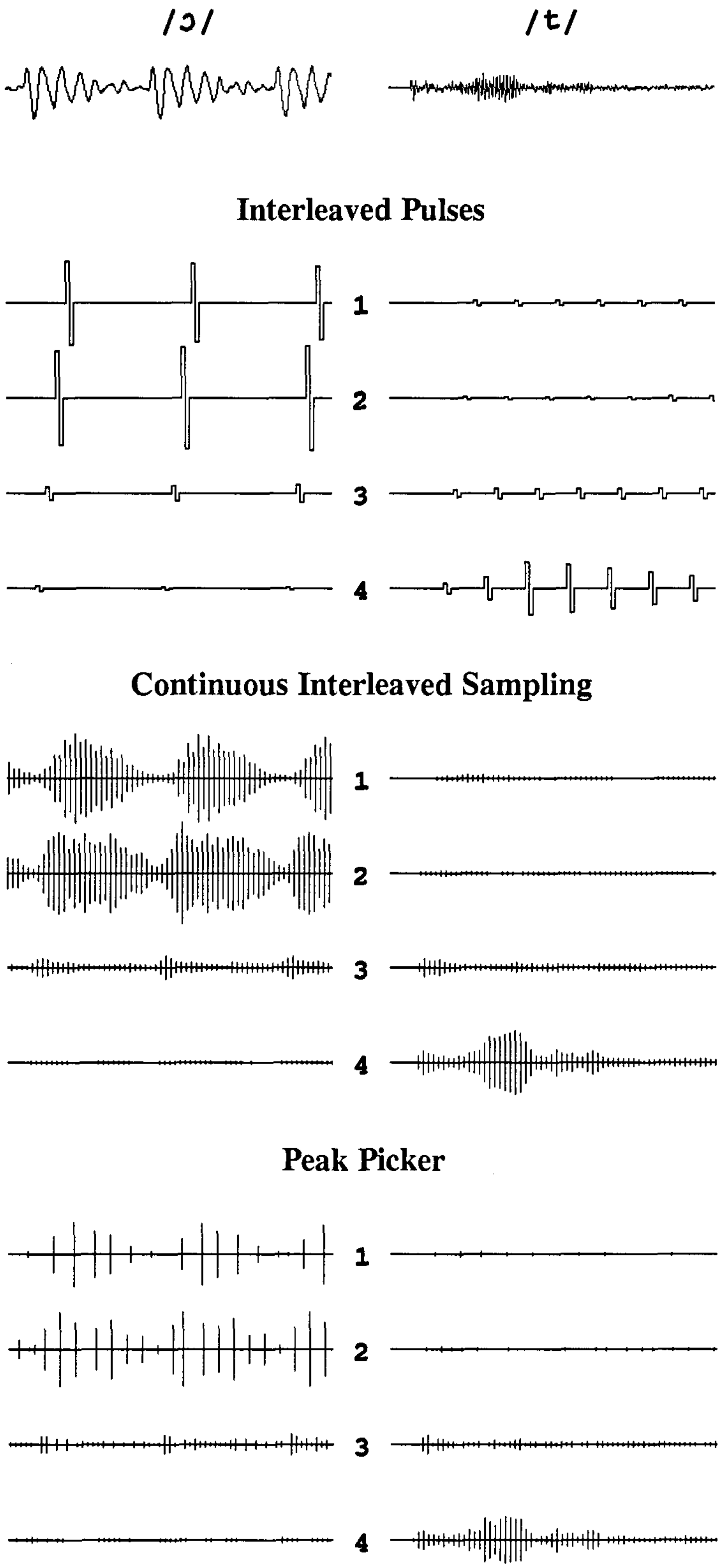


Figure 8. Waveforms of three types of pulsatile processors. The duration of each trace is 25.4 ms. Note that the prolonged stimulation in the IP processor for the /t/ burst is a consequence of the long time constant of the lowpass filters in the envelope detectors (25 Hz cutoff versus 400 Hz cutoff for the other processors).

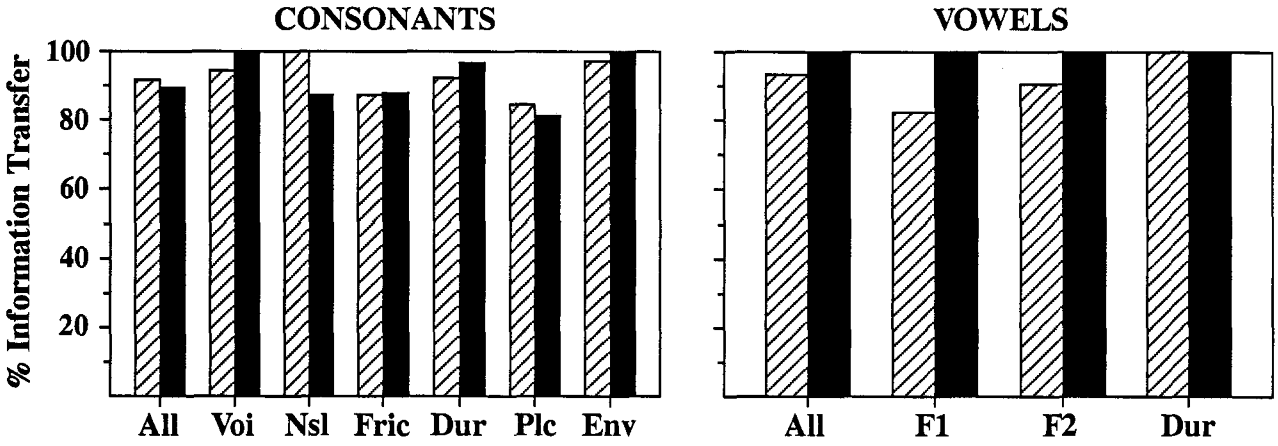


Figure 9. Comparison of speech test scores for an early implementation of a CIS processor (striped bars) and a peak picker (PP) processor (solid bars) evaluated at the same time. Twenty presentations of each of 16 consonants were used in the consonant identification test for both processors. The presentations were equally divided between the male and female speakers. For the vowel identification tests 18 presentations were used for evaluation of the CIS processor (striped bars) and 12 were used for the PP processor (solid bars). The presentations were equally divided between the male and female speakers.

SR2 remarked that percepts produced with the PP processor were more "pitch appropriate" than percepts produced with the contemporaneous versions of the CIS processor, particularly for low frequency sounds such as the fundamental frequency of voiced speech. While the PP processor sounded a bit more natural to SR2, both processors were judged by him to be highly intelligible, with no clear difference between processors in conveying connected speech. (SR2 has scored substantially higher with subsequently developed CIS versions.)

Hybrid PP/CIS Processor

In more recent studies we have evaluated a hybrid of the PP and CIS strategies. In this PP/CIS processor PP stimuli were delivered to the apical two electrodes in the Ineraid array and CIS stimuli were delivered to the remaining four electrodes. The speech processor was programmed to examine the signals from the envelope detectors in the apical-most two channels just before the scheduled delivery of a CIS pulse on one of the more basal channels. If the processor detected a peak in one or both of the apical channels, then the CIS pulse would be delayed to allow the delivery of a (nonsimultaneous) PP pulse to each channel with a detected peak. This process was repeated for each CIS pulse.

Results from an initial evaluation of this hybrid PP/CIS processor are presented in Fig. 10. In addition,

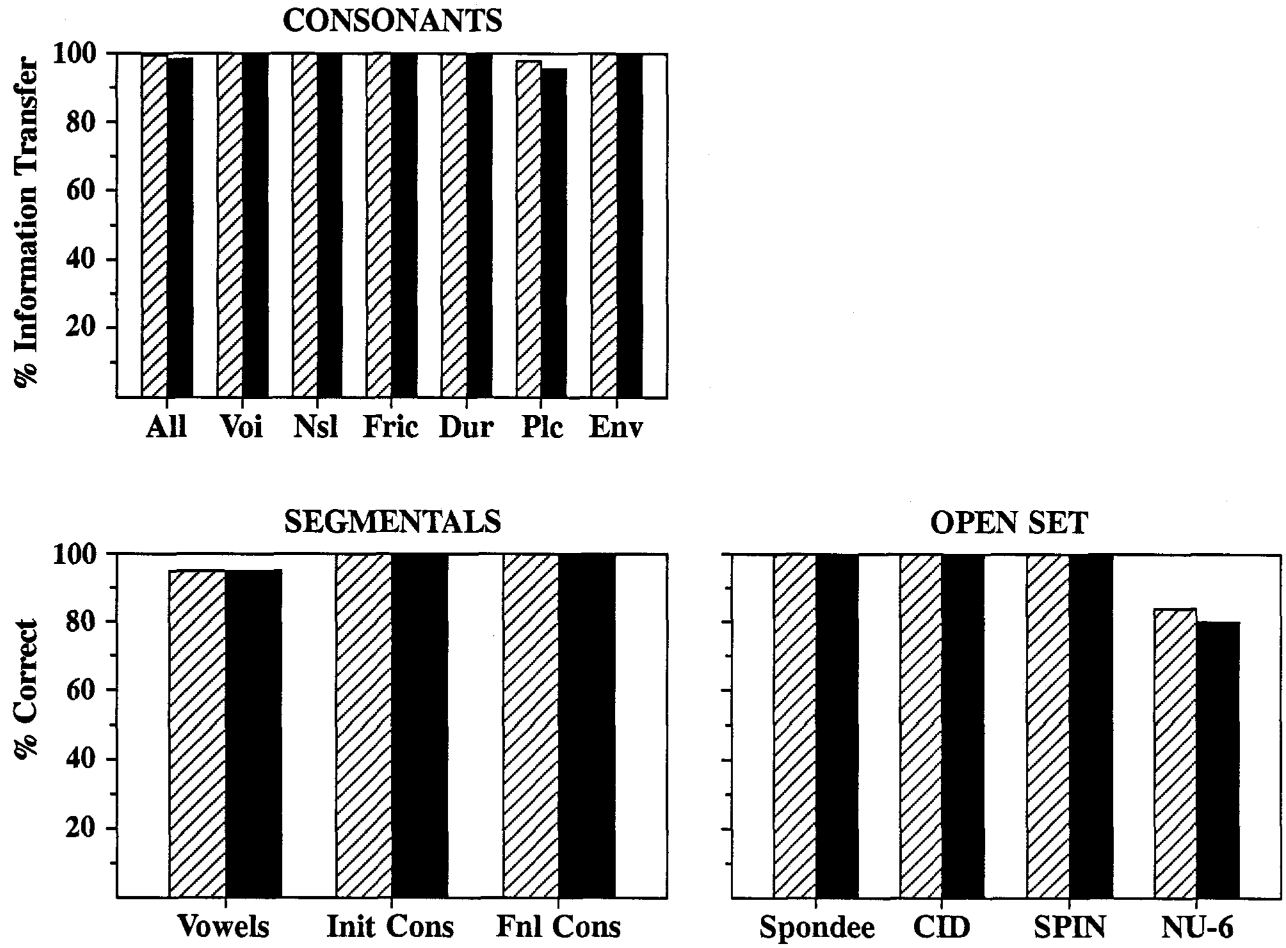


Figure 10. Comparison of speech test scores for a standard CIS processor (striped bars) and a hybrid PP/CIS processor (solid bars). Five presentations of each of 24 consonants by the male speaker were used in the consonant identification tests for both processors. Tests of vowel identification were not included in the comparison of these two processors.

results from a CIS processor with parameters similar or identical to those of the CIS channels in the PP/CIS processor are shown. The consonant test used for these evaluations included 24 consonants. Tests were conducted with the male speaker only. The CID, SPIN and NU-6 tests were conducted using novel lists of recorded sentences and words (with a different male speaker).

Clearly, both processors support high levels of speech recognition. All scores except those for the NU-6 test approach or hit the 100% ceiling. The NU-6 scores are 84 and 80% correct for the CIS and PP/CIS strategies, respectively.

Anecdotally, the PP/CIS processor sounded quite natural, especially for music. Indeed, SR2, who was a musician before he lost his hearing, indicated that percepts produced with the PP/CIS processor had greater "pitch appropriateness" and musical clarity than percepts produced with his clinical CA processor or with the CIS processor (this observation was made while listening to small ensemble jazz recordings, which were familiar to the subject only through use of his clinical CA processor). Because the sound was so enjoyable to SR2 ("music is wonderful through this [PP/CIS] processor, like nothing I've even dreamed of ever hearing again with my implant"), we took some time off from testing so that he could listen to tapes of music he remembered from years ago (rock music familiar to the subject before loss of his normal hearing, which he had chosen not to listen to with his clinical CA device). Again, SR2 heard nuances in the material that could not be perceived with the other processors.

These high levels of performance for the PP/CIS strategy are encouraging. The preliminary results, along with anecdotal comments, suggest that use of PP stimuli, at least for the apical channels, may provide an improved representation of frequencies in the F0 and F1 ranges, at least for some of the better subjects.

V. Auditory Brainstem Implant

The Auditory Brainstem Implant (ABI) has been used to restore some hearing for people with bilateral loss of cochlear nerve continuity. To date, approximately 20 people have been implanted with the ABI device, in association with removal of the bilateral acoustic neuromas found in patients with neurofibromatosis type II.

We have studied two of these patients, in collaboration with Robert V. Shannon and others at the House Ear Institute. The studies were conducted in our laboratory at Duke University Medical Center, beginning in the fall of 1989.

The ABI was placed in the first patient during removal of his second acoustic tumor. In contrast, the device was placed in the second patient during removal of the first of her bilateral tumors. Thus the second patient still had normal hearing in one ear at the time of our tests and no experience with prosthetic stimulation of her implant. The first patient was totally deaf without his prosthesis, and had approximately five months of experience with his ABI at the time of our tests. Both subjects had percutaneous access to their implanted electrodes, and in both cases only one of the two implanted electrodes offered the possibility of stimulating purely auditory percepts.

Results for the first subject are shown in Fig. 11. A single-channel *continuous sampling* (CS) processor was compared with the subject's clinical HEI processor (identical in most respects to the 3M/House processor). The stimuli presented by the CS processor, a single-channel variation of CIS processors, consisted of a train of short duration pulses whose amplitudes were modulated (via a logarithmic mapping function) with the envelope of the broadband speech signal. The tests included identification of 16 consonants, using male and female speakers; identification of 8 vowels, using male and female speakers; the segmental tests of the Minimal Auditory Capabilities (MAC) battery (Owens et al., 1985); and all open-set tests of the MAC battery except for the SPIN test, which was omitted for this subject. All tests were conducted with hearing alone, using single presentations of recorded material and with no feedback as to correct or incorrect responses.

As is obvious from the figure, use of the CS processor produced large gains in the transmission of consonant information. In particular, scores for the temporal features of voicing, frication, duration, and envelope cues are much higher with the CS processor. In addition, the score for place of articulation is more than doubled with the application of the CS processor. The only score not improved with the CS processor is the one for nasality, which is about the same for the two processors.

Transmission of vowel features is about the same for the two processors. Also, the scores for the vowel test in the MAC battery are essentially equivalent for the two processors.

In contrast to the vowel scores, remarkable gains in open-set recognition are produced with the use of the CA processor. The score for spondee recognition is increased from 2 to 40% correct, for CID sentences from 11 to 25% correct, and for NU-6 words from 2 to 12% correct.

These increases, particularly for open-set recognition, are all the more remarkable when one considers

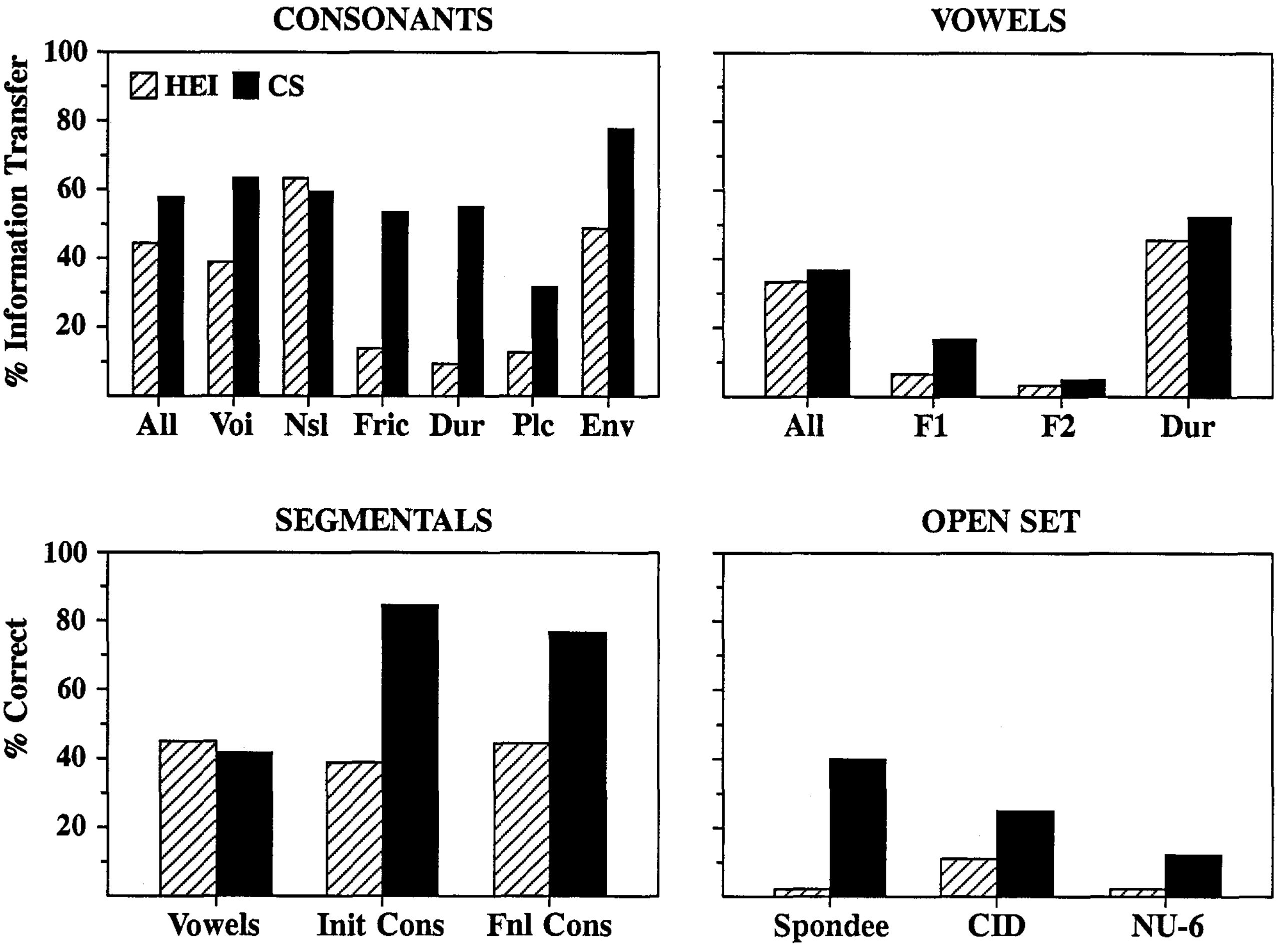


Figure 11. Comparison of speech test scores for the first ABI patient. Scores for the HEI processor are indicated by the striped bars, and those for the CS processor by the solid bars. The top panels show relative information transfer for articulatory and acoustic features of consonants and vowels (see Miller and Nicely, 1955). The features for consonants include overall transmission (All), voicing (Voi), nasality (Nsl), frication (Fric), duration (Dur), place of articulation (Plc), and envelope cues (Env). The features for vowels include overall transmission (All), first formant frequency (F1), second formant frequency (F2), and duration (Dur). Twenty presentations of each of 16 consonants were used in the consonant identification tests for both processors, and eighteen presentations of each of 8 vowels were used in the vowel identification tests for both processors. Presentations for both the consonant and vowel tests were equally divided between male and female speakers. The bottom panels show scores from the segmental and open-set tests of the Minimal Auditory Capabilities (MAC) battery. The CS processor (processor SS2B) used 110 μs/phase pulses, presented at the rate of 1818 pps. The cutoff frequency of the lowpass filter in the envelope detector was 400 Hz.

the disparity in experience with the two processors. This subject had five months of daily experience with his HEI processor, but only several hours of (aggregated) experience with CS processors before these tests were conducted.

Studies with the second subject were complicated by the fact that she had normal hearing, and that she lacked any experience with electrical stimulation.

Most studies with her were directed at acclimating her to electrically evoked percepts and to initial evaluations of the CS strategy as an adjunct to lipreading. As indicated in detail in QPR 6 for this project (in the section on "Parametric Variations and the Fitting of Speech Processors for Single-Channel Brainstem Prostheses"), use of the CS strategy in conjunction with lipreading (from the Iowa laser videodisc images) produced consonant identification scores in the high 90s. Such scores are compatible with high levels of open-set speech recognition. Thus, even in a totally naive listener, the CS strategy demonstrated its potential as an adjunct to lipreading.

While these findings are most encouraging, recent results from studies with CIS processors suggest that substantial improvements in speech recognition might be obtained with additional channels. In particular, consonant identification increased almost linearly with increases in channel number from 1 to 6 for a subject using a scala tympani implant (Lawson et al., 1992). Effective use of such additional channels for the ABI device would of course depend on the number of perceptually distinct channels of stimulation.

The present HEI Implant has two large electrode surfaces that overlie the dorsal cochlear nucleus. In most cases, only one of these electrodes is useful, in that (monopolar) stimulation of the other produces various nonauditory percepts such as dizziness. In the few cases in which both electrodes produce auditory sensations, the percepts have been described as identical (Shannon, personal communication).

Although distinct auditory percepts have not been demonstrated in ABI patients, studies of Frederickson and Gerken (1977) indicate that penetrating electrodes, properly positioned (in the ventral cochlear nucleus), can produce tonotopically restricted patterns of activation in the central auditory system. Use of such electrodes may allow the effective application of multichannel CIS processors.

Electrodes under development include the penetrating electrodes at the University of Michigan and at HEI/Huntington. In addition, Cochlear Corporation has developed an array of surface electrodes (including 8 contacts) in a cooperative effort with HEI. We plan to continue our collaborative studies with Bob Shannon and others at HEI to (a) study additional patients with the present electrode system and (b) study patients who might be implanted in the future with one of the new electrode systems.

VI. Record of Reporting Activity for NIH Project N01-DC-9-2401

We have maintained a high level of reporting activity throughout the project. This activity, for the period of May 1, 1989 through July 31, 1992, includes publication of 10 papers, presentation of 16 invited lectures, presentation of 5 other talks or posters, and publication of 8 abstracts. In many cases all expenses were reimbursed for the invited lectures, allowing us to present project results at no cost to the project.

We also note that members of the RTI team have served as the chair or moderator of sessions at international meetings, and that Blake Wilson served as the General Chair for the *1991 Conference on Implantable Auditory Prostheses*, held in Pacific Grove, CA, June 2-7.

The citations are:

Papers

Wilson BS, Finley CC, Lawson DT, Wolford RD, Zerbi M (1993). Design and evaluation of a continuous interleaved sampling (CIS) processing strategy for multichannel cochlear implants. *J Rehab Res Devel*, in press.

Wilson BS, Lawson DT, Finley CC, Wolford RD (1992). Importance of patient and processor variables in determining outcomes with cochlear implants. *J Speech Hear Res*, in press.

Lawson DT, Wilson BS, Finley CC (1992). New processing strategies for multichannel cochlear prostheses. *Prog Brain Res*, in press.

Wilson BS (1992). Signal processing. In R. Tyler (Ed.), *Cochlear Implants: Audiological Foundations*, Singular Publishing Group, San Diego, CA, pp. 31-81, in press.

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, Finley CC, Wolford RD (1991). Coding strategies for multichannel cochlear prostheses. *Am J Otol* 12, Suppl. 1: 56-61.

Finley CC, Wilson BS, White MW (1990). Models of neural responsiveness to electrical stimulation. In J.M. Miller and F.A. Spelman (Eds.), *Cochlear Implants: Models of the Electrically Stimulated Ear*, Springer-Verlag, New York, pp. 55-96.

Wilson BS, Finley CC, Lawson DT (1990). Representations of speech features with cochlear implants. In J.M. Miller and F.A. Spelman (Eds.), *Cochlear Implants: Models of the Electrically Stimulated Ear*, Springer-Verlag, New York, pp. 339-376.

Finley CC (1990). Radial bipolar electrode placement in scala tympani: Effects on neural potential profiles and longitudinal spread of excitation. In *Proc Twelfth Ann Conf Engineering in Medicine and Biology Soc*, IEEE Press, New York, pp. 2290-2291.

Finley CC (1989). A finite-element model of bipolar field patterns in the electrically stimulated cochlea -- Two and three dimensional approximations and tissue parameter sensitivities. In *Proc Eleventh Ann Conf Engineering in Medicine and Biology Soc*, IEEE Press, New York, pp. 1059-1060.

Chaired Meetings and Sessions

Wilson BS (1991). General Chair. *1991 Conference on Implantable Auditory Prostheses*, Pacific Grove, CA, June 2-7.

Wilson BS (1990). Moderator, Session on Speech Processing. *Second International Cochlear Implant Symposium*, Iowa City, IA, June 4-8.

Finley CC (1989). Chair, Session on Cochlear Prostheses. *Eleventh Ann. Conf. Engineering in Medicine and Biology Soc.*, Seattle, WA, November 8-12.

Wilson BS (1989). Steering Committee Member, *Engineering Foundation Conference on Implantable Auditory Prostheses*, Potosi, MO, July 30 to August 4.

Invited Presentations

Wilson BS (1989, 1990 and 1992). Speech processors for auditory prostheses. *Neural Prosthesis Workshop*, National Institutes of Health, Bethesda, MD, October.

Wilson BS (1992). Processing strategies for multichannel cochlear implants. Faculty lecture, *Fourth Symposium: Cochlear Implants in Children*, Kansas City, MO, February 14 and 15. [Abstract published in Symposium Proceedings]

Wilson BS, Lawson DT, Finley CC (1991). A new processing strategy for multichannel cochlear prostheses. *International Symposium on Natural and Artificial Nervous Control of Hearing and Balance*, Rheinfelden, Switzerland, September 4-8. [Presented by DT Lawson; abstract published in Symposium Proceedings]

Wilson BS (1991). A new coding strategy for cochlear implants. *American Neurotology Society*, Kansas City, MO, September 21.

Wilson BS (1991). New levels of speech recognition with cochlear implants. *1991 Conference on Implantable Auditory Prostheses*, Pacific Grove, CA, June 2-7.

Finley CC (1991). Models of potential distributions for various types and placements of electrodes. *1991 Conference on Implantable Auditory Prostheses*, Pacific Grove, CA, June 2-7.

Wilson BS (1991). Strategies for representing speech with cochlear implants. In the special session on speech perception and hearing handicap, *Meeting of the Acoustical Society of America*, Baltimore, MD, April 29 to May 3. [Abstract published in *J Acoust Soc Am* 89, Suppl. 1, p. 1957]

Shannon RV (moderator), Wilson BS, Eddington DK, Walliker J, Pfingst BE, Patrick JF, Rosen S (panelists) (1990). Round table discussion on "Future directions in Speech Processing." *Second International Cochlear Implant Symposium*, Iowa City, IA, June 4-8.

Wilson BS (1990). Design of cochlear prostheses. Presented in the special session on "Cochlear Implants in Children," *AAAS Meeting*, New Orleans, LA, February 15-20.

Wilson BS (1990). Processing strategies for cochlear implants. Faculty lecture, *Third Symposium on Cochlear Implants in Children*, Indianapolis, IN, January 26 and 27. [Abstract published in Symposium Proceedings]

Farmer JC Jr., Javel E, McElveen JT Jr., Wilson BS (1989). Advances in cochlear implants. *Surgical Grand Rounds*, Duke University Medical Center, December 13.

Finley CC (1989). Cochlear implants. Presented at the House Ear Institute, Los Angeles, CA, November 15.

Wilson BS (1989). Comparison of analog and pulsatile coding strategies for multichannel cochlear prostheses. Presented at the University of Iowa, Iowa City, IA, August 28.

Wilson BS (1989). Within patient evaluation of speech processors. *Engineering Foundation Conference on Implantable Auditory Prostheses*, Potosi, MO, July 30 to August 4.

Additional Presentations and Abstracts

Wilson BS, Finley CC, Lawson DT, Wolford RD (1991). New levels of speech recognition with cochlear implants. *1991 Midwinter Meeting of the Association for Research in Otolaryngology*, St. Petersburg, FL, February 3-7. [Abstract published in *ARO Abstracts*, 14th Midwinter Research Conference, p. 35]

Finley CC (1991). Bipolar electrode placement in scala tympani: Effects on neural potential profiles, longitudinal recruitment and activating functions. *1991 Midwinter Meeting of the Association for Research in Otolaryngology*, St. Petersburg, FL, February 3-7. [Abstract published in *ARO Abstracts*, 14th Midwinter Research Conference, p. 52]

Wilson BS, Finley CC, Lawson DT, Wolford RD (1990). A new processing strategy for multichannel cochlear implants. *Second International Cochlear Implant Symposium*, Iowa City, IA, June 4-8. [Abstract published in Symposium Proceedings]

Finley CC, Wilson BS (1990). Spiral ganglion cell body effects on neural response latency in the electrically stimulated cochlea. *1990 Midwinter Meeting of the Association for Research in Otolaryngology*, St. Petersburg, FL, February 4-8. [Abstract published in *ARO Abstracts*, 13th Midwinter Research Conference, pp. 331-332]

Finley CC (1989). Electric field patterns produced by intracochlear stimulation. Poster presentation, *Engineering Foundation Conference on Implantable Auditory Prostheses*, Potosi, MO, July 30 to August 4.

VII. Suggestions for Future Research

Large gains in open-set speech recognition have been demonstrated in this project for users of cochlear and auditory brainstem prostheses. While the present results are most encouraging, many important questions remain unanswered, and many possibilities for further improvement remain unexplored. These questions and possibilities are being addressed in our laboratory, under the auspices of a new project with the Neural Prosthesis Program, and elsewhere. Particularly promising lines of investigation include the following:

- Systematic evaluation of parameter choices for CIS processors, including detailed evaluation of tradeoffs among pulse duration, pulse rate, interval between sequential pulses, and cutoff frequency of the lowpass filters in the envelope detectors.
- Systematic evaluation of how those choices vary across subjects.
- Identification of mechanisms underlying patient variability (see QPR 8 and Wilson et al., 1992).
- Application of CIS strategies in the MiniMed and other devices, that would allow implementations with more than six channels and with electrode coupling arrangements other than the monopolar configuration of the Ineraid device.
- Evaluation of possible learning effects with extended use of CIS processors.
- Continued development of highly-flexible portable processors to support such field studies with CIS and other strategies.
- Design and application of new implant devices capable of presenting stimuli at high rates to a large number of electrodes, with a variety of externally controlled coupling configurations.
- Development and evaluation of techniques to reduce the deleterious effects of environmental noise on processor performance.
- Continued studies, with additional subjects and tests, to evaluate further the PP and PP/CIS strategies.
- Continued development of new processing strategies, which may produce further increases in performance.
- Continued studies with the Auditory Brainstem Implant, especially if patients implanted with any of the new multichannel penetrating electrodes become available for such studies.
- Development of new test materials, to provide sensitive measures of performance for subjects who now, with certain speech processors and implant devices, encounter ceiling effects with all of our standard tests.

VIII. Acknowledgements

We thank the subjects of the described studies for their enthusiastic participation. We also are pleased to acknowledge the important scientific contributions of many colleagues, including, for the studies reviewed here, Michael F. Dorman, Donald K. Eddington, Albert Maltan, William M. Rabinowitz, Robert V. Shannon, Robert D. Wolford and John Wygonski.

IX. References

Bess FH, Townsend TH (1977). Word discrimination for listeners with flat sensorineural hearing losses. *J Speech Hear Disorders* 42: 232-237.

Clark GM (1987). The University of Melbourne-Nucleus multi-electrode cochlear implant. *Adv Oto-Rhino-Laryngol* 38: 1-189.

Dorman MF, Soli S, Dankowski K, Smith LM, McCandless G, Parkin J (1990). Acoustic cues for consonant identification by patients who use the Ineraid cochlear implant. *J Acoust Soc Am* 88: 2074-2079.

Dowell RC, Brown AM, Mecklenburg DJ (1990). Clinical assessment of implanted deaf adults. In *Cochlear Prostheses*, G.M. Clark, Y.C. Tong and J.F. Patrick (Eds.), 193-205. Edinburgh: Churchill Livingstone.

Dowell RC, Seligman PM, Blamey PJ, Clark GM (1987). Evaluation of a two-formant speech-processing strategy for a multichannel cochlear prosthesis. *Ann Otol Rhinol Laryngol* 96 (Suppl. 128): 132-134.

Dubno JR, Dirks DD (1982). Evaluation of hearing-impaired listeners using a nonsense syllable test. I. Test reliability. *J Speech Hear Res* 25: 135-141.

Eddington DK (1980). Speech discrimination in deaf subjects with cochlear implants. *J Acoust Soc Am* 68: 885-891.

Eddington DK (1983). Speech recognition in deaf subjects with multichannel intracochlear electrodes. *Ann NY Acad Sci* 405: 241-258.

Frederickson CJ, Gerken GM (1977). Masking of electrical by acoustic stimuli: Behavioral evidence for tonotopic organization. *Science* 198: 1276-1278.

Gantz BJ, McCabe BF, Tyler RS, Preece JP (1987). Evaluation of four cochlear implant designs. *Ann Otol Rhinol Laryngol* 96, Suppl. 128: 145-147.

Hall JW, Haggard MP, Fernandes MA (1984). Detection in noise by spectro-temporal pattern analysis. *J Acoust Soc Am* 76: 50-56.

Lawson DT, Wilson BS, Finley CC (1992). New processing strategies for multichannel cochlear prostheses. *Prog Brain Res*, in press.

Merzenich MM, Rebscher SJ, Loeb GE, Byers CL, Schindler RA (1984). The UCSF cochlear implant project. State of development. *Adv Audiol* 2: 119-144.

Miller, GA, Nicely, PE (1955). An analysis of perceptual confusions among some English consonants. *J Acoust Soc Am* 27: 338-352.

Moore BCJ (1992). Comodulation masking release and across-channel masking. In *Audition, Speech and Language*, M.E.H. Schouten (Ed.). Berlin: Mouton de Gruyter, in press.

Owens E, Kessler DK, Raggio M, Schubert ED (1985). Analysis and revision of the Minimal Auditory Capabilities (MAC) battery. *Ear Hear* 6: 280-287.

Pfingst BE (1984). Operating ranges and intensity psychophysics for cochlear implants. *Arch Otolaryngol* 110: 140-144.

Rabiner LR, Shafer RW (1978). *Digital Processing of Speech Signals.* Englewood Cliffs, NJ: Prentice-Hall.

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

Shannon RV (1992). Psychophysics. In *Cochlear Implants: Audiological Foundations*, R.S. Tyler (Ed.), Chapter 3. San Diego, CA: Singular Publishing Group, in press.

Tye-Murray N, Tyler RS (1989). Auditory consonant and word recognition skills of cochlear implant users. *Ear Hear* 10: 292-298.

Tyler RS, Preece JP, Lansing CR, Otto SR, Gantz BJ (1986). Previous experience as a confounding factor in comparing cochlear-implant processing schemes. *J Speech Hear Res* 29:282-287.

Tyler RS, Preece JP, Lowder MW (1987). *The Iowa audiovisual speech perception laser videodisc.* Laser Videodisc and Laboratory Report, Department of Otolaryngology -- Head and Neck Surgery, University of Iowa Hospitals and Clinics, Iowa City, Iowa.

Tyler RS, Tye-Murray N (1991). Cochlear implant signal-processing strategies and patient perception of speech and environmental sounds. In *Cochlear Implants: A Practical Guide*, H. Cooper (Ed.), 58-83. San Diego, CA: Singular Publishing Group.

White MW, Merzenich MM, Gardi JN (1984). Multichannel cochlear implants: Channel interactions and processor design. *Arch Otolaryngol* 110: 493-501.

Wilson BS (1992). Signal processing. In *Cochlear Implants: Audiological Foundations*, R.S. Tyler (Ed.), Chapter 2. San Diego, CA: Singular Publishing Group, in press.

Wilson BS, Finley CC, Lawson DT (1990a). Representations of speech features with cochlear implants. In *Cochlear Implants: Models of the Electrically Stimulated Ear*, J.M. Miller and F.A. Spelman (Eds.), 339-376. New York: Springer-Verlag.

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

Wilson BS, Finley CC, Lawson DT, Wolford RD, Zerbi M (1993). Design and evaluation of a continuous interleaved sampling (CIS) processing strategy for multichannel cochlear implants. *J Rehab Res Devel*, in press.

Wilson BS, Lawson DT, Finley CC (1990b). Speech processors for auditory prostheses. Fourth Quarterly Progress Report, NIH project N01-DC-9-2401.

Wilson BS, Lawson DT, Finley CC, Wolford RD (1991b). Coding strategies for multichannel cochlear prostheses. *Am J Otol* 12, Suppl. 1: 56-61.

Wilson BS, Lawson DT, Finley CC, Wolford RD (1992). Importance of patient and processor variables in determining outcomes with cochlear implants. *J Speech Hear Res*, in press.

Wilson BS, Lawson DT, Finley CC, Zerbi M (1991c). Speech processors for auditory prostheses. Ninth Quarterly Progress Report, NIH project N01-DC-9-2401.