Final Report

September 26, 1985 through March 31, 1989

NIH project N01-NS-5-2396

"Speech Processors for Auditory Prostheses"

Prepared by

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

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

This QPR is being sent to you before it has been reviewed by the staff of the Neural Prosthesis Program

CONTENTS

I. Introduction

The purpose of this project is to design and evaluate speech processors for implantable auditory prostheses. Ideally, the processors will extract (or preserve) from speech those parameters that are essential for intelligibility and then appropriately encode these parameters for electrical stimulation of the auditory nerve or central auditory structures. Major activities and accomplishments over the period of the project -- from late September, 1985 to late March, 1989 -- include the following:

1. Investigation of psychophysical performance and speech perception in tests with subjects implanted with various types of cochlear prostheses, including the experimental UCSF prosthesis with a percutaneous cable, the 4-channel UCSF/Storz prosthesis, the single-channel extracochlear version of the 3M/Vienna prosthesis, the experimental multichannel prosthesis developed at 3M, the Nucleus prosthesis, and the Symbion prosthesis;
2. Development and evaluation of several new processing strategies for cochlear prostheses, including an interleaved pulses (IP) processor for multichannel prostheses, a hybrid compressed analog (CA)/IP processor for multichannel prostheses, and a two channel processor based on a strategy first described by Breeuwer and Plomp;
3. Direct comparisons of these new processors with each other and with the standard processors of clinical devices using within subject designs;
4. Investigation of differences in loudness and pitch perception under conditions of stimulation with monopolar and radial bipolar electrodes;
5. Development of software and laboratory equipment to support the above psychophysical and speech perception studies;
6. Development and application of portable, real-time processors for implementing various multichannel coding strategies, including the IP and Breeuwer/Plomp (B/P) strategies mentioned above;
7. Further development of models of neural responsiveness to intracochlear electrical stimulation;
8. Collaboration with the UCSF team in the development of the speech processor, transcutaneous transmission system (TTS), and clinical fitting system for a next-generation cochlear prosthesis;
9. Publication of 10 papers and 15 abstracts; and
10. Further dissemination of project results in 43 presentations (36 invited) at national and international meetings.

In this report we describe the activities indicated in points 1-4 and 8 above. A summary is presented in a section of concluding remarks. This last section also briefly indicates our plans and recommendations for future research.

Additional information on all project activities may be found in our quarterly progress reports and publications. The three appendices to this report provide access to this information. Appendix 1 presents a table of key contents of the quarterly reports; Appendix 2 lists the publications and presentations for the project; and Appendix 3 presents reprints of the publications.

II. Direct Comparisons of Analog and Pulsatile Coding Strategies

The development and application of cochlear prostheses have improved the quality of life for many deaf individuals. Much work remains to be done, however, in order to achieve high levels of speech recognition in a majority of patients. In particular, we do not fully understand how different speech processor and implanted electrode designs affect the perception of speech and other sounds. Also, we have only fragmentary and primitive knowledge of how various differences among patients affect their outcomes.

The principal characteristics of five different cochlear implant systems are presented in Table 1. These characteristics include the type of electrode used, the number of stimulation channels, and the method for processing speech inputs. As is evident from the table, a tremendous diversity exists in current approaches to prosthesis design. Further details on the systems listed in Table 1 can be found in recent reviews [Gantz, 1987; Millar et al., 1984; Moore, 1985; Parkins, 1986; Pfingst, 1986] and in primary sources [Clark et al., 1987; Danley and Fretz, 1982; Eddington, 1983; Hochmair and Hochmair-Desoyer, 1985; Merzenich, 1985].

A remarkable finding from evaluations of the disparate approaches to prosthesis design is that several of these approaches can produce high levels of speech perception in some patients. Unfortunately, the converse is also true in that poor levels of performance are found for other patients using the same approaches.

These facts are illustrated in Table 2, which compares the average values and standard deviations of speech recognition scores for the three multichannel systems listed in Table 1. All the data presented in Table 2 were obtained from studies in which single presentations of recorded material were used. The data reported for the Nucleus device are those for the new "F0/F1/F2" processing strategy. Note that the standard deviations are large for all measures and devices. This variability reflects the observation that, for any of the listed devices, excellent performance is found for some patients, while poor performance is found for others. Finally, it is worth noting that experience with an implant can have a profound effect on the scores for these (and other) tests. With the Storz device, for example, both the NU6 word and CID sentence scores show significant increases ($p < 0.02$ for the NU6 test and $p < 0.01$ for the CID test) from the 3 month to the 12 month studies.

Much of this heterogeneity of outcomes may be attributed to variation among patients. The importance of learning effects associated with experience has just been mentioned. Other potentially important variables include differences among patients in the survival of neural elements in the implanted ear, integrity of the central auditory pathways, and cognitive and language skills. The likely influence of patient variables on outcome obviously complicates comparisons aimed at identifying the best types of electrode and speech processor designs for cochlear prostheses. Indeed, such comparisons are not meaningful when small populations of subjects are used for the evaluation of each system or when test procedures used at the various laboratories involved in the evaluation of individual systems are different.

Table 1. Cochlear implant systems.

Device	Electrode(s)	Stimulus Channels/ Stimulus Sites	Processing Strategy
3M/House	monopolar	1 / 1	amplitude modulation of 16 kHz carrier
3M/Vienna	longitudinal bipolar or extracochlear	1 / 1	compressed analog with frequency equalization
Nucleus	longitudinal bipolar	2 / 21	feature extraction (F0, F1 and F2), pulsatile stimulation
Symbion	monopolar	4 / 4	compressed analog with filtering
Storz	radial bipolar	4 / 4	compressed analog with filtering

Table 2. Average percent correct scores for recognition of NU6 words and CID sentences with multichannel cochlear prostheses. The standard deviations of these scores are given in parentheses.

Device	NU6	CID	N	Months Postimplant	Citation
Nucleus	6.5 (6.4)	20.7 (19.5)	53/49*	3	Brimacombe et al., 1988
Symbion	8.5 (10.9)		11	>7	Gantz, 1987
Storz	7.6 (8.2)	21.0 (16.5)	14	3	Wilson et al., 1988d
	20.2 (15.0)	45.9 (26.8)	12	12	Wilson et al., 1988d

*N = 53 for the NU6 test and N = 49 for the CID test.

Recognition of the problem just described helped to initiate two major studies in which relatively large populations of patients implanted with different devices are being tested in a uniform and consistent manner. One of these studies is a cooperative effort among VA medical centers and the other study is being conducted at the University of Iowa [Gantz et al., 1988]. Results from these studies should be of great value in establishing expected levels of performance for contemporary cochlear prostheses.

A complementary method for the comparison of prosthesis systems is to evaluate different processing strategies and electrode coupling configurations in tests with the same implant patient. A key advantage of such tests is that controls are provided for patient variables. Thus, for a given patient with a fixed pattern of nerve survival and fixed levels of cognitive skill, etc., the performance levels of different prosthesis systems can be compared directly.

In studies conducted in collaboration with investigators at the University of California at San Francisco (UCSF) and at Duke University Medical Center, our team compared a variety of speech processing strategies in tests with patients implanted with the UCSF/Storz electrode array [Wilson et al., 1988a,b,c,d]. Some of the largest differences in performance among processing strategies were found in comparisons between the compressed analog (CA) processor of the present UCSF/Storz prosthesis and a type of "interleaved pulses" (IP) processor which delivers pulses in sequence to the different channels in the implanted electrode array. To show how the design of the processor can affect the outcome for individual patients, we will restrict ourselves here to brief descriptions of tests related to these two types of processor. Detailed descriptions of these tests, along with the results obtained with other processing strategies, are presented elsewhere [Wilson et al., 1985, 1986a, 1987b, 1988a,b,c; Wilson et al., in press].

Processing Strategies

In the clinical UCSF/Storz device alternate pairs of the 16 available electrodes are stimulated simultaneously with the CA outputs of a four-channel speech processor. The basic functions of this processor are to compress the wide dynamic range of input speech signals into the narrow dynamic range available for electrical stimulation and then to filter the compressed signal into individual frequency bands for presentation to each pair of stimulated electrodes. Typical waveforms of the CA processor are shown in Fig. 1. The top trace in each panel is the input signal, which in this case is the word "bought." The other waveforms in each panel are the filtered output signals for 4 channels of intracochlear stimulation. The bottom left panel shows an expanded display of waveforms during the initial part of the vowel in "bought," and the bottom right panel shows an expanded display of waveforms during the final /t/. The lower panels in Fig. 1 thus exemplify differences in waveforms for voiced and unvoiced intervals of speech.

In the voiced interval the relatively large outputs of channels 1 and 2 reflect the low-frequency formant content of the vowel, and in the unvoiced interval the relatively large outputs of channels 3 and 4 reflect the high-frequency noise content of the /t/. In addition, the clear periodicity in the

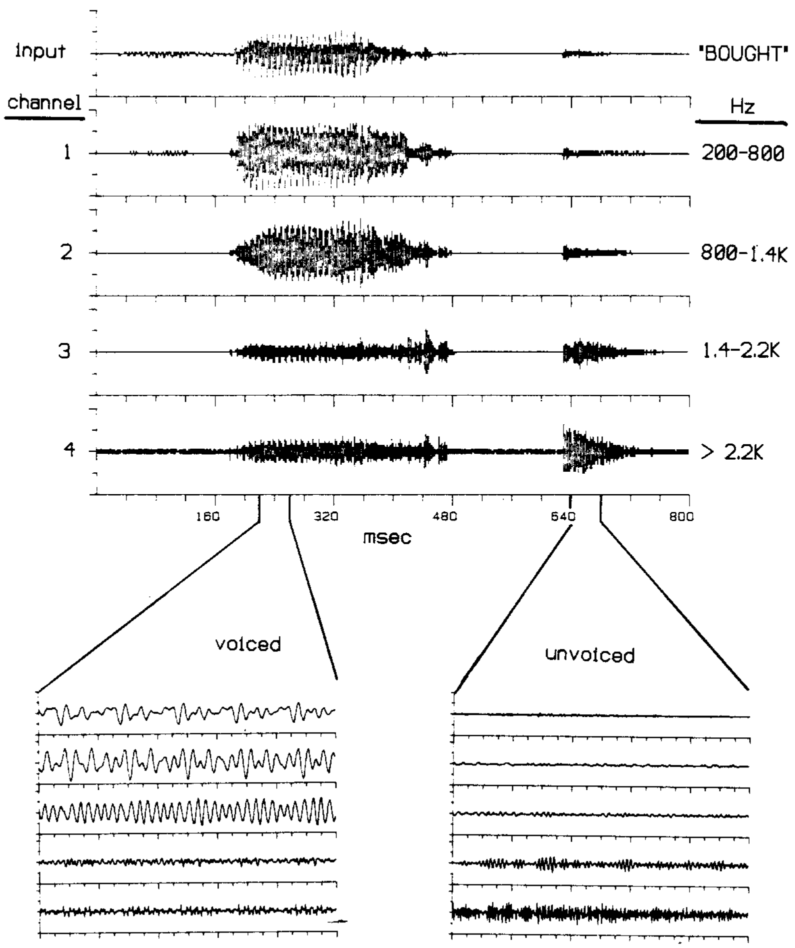


Fig. 1. Waveforms of a compressed analog (CA) processor.

waveforms of channels 1 and 2 reflects the fundamental and first formant frequencies of the vowel during the voiced interval, and the lack of periodicity in the output of any channel reflects the noise-like quality of the /t/ during the unvoiced interval. As has been described elsewhere [Schindler and Kessler, 1987; Schindler et al., 1986, 1987; Wilson et al., 1988d], this representation of speech features can support high levels of open-set recognition for many (but not all) of the patients implanted with the UCSF/Storz prosthesis.

A concern associated with the use of multichannel CA processors is that of channel interactions [White et al., 1984]. Simultaneous stimulation of two or more channels with continuous waveforms results in summation of the electrical fields from the individual bipolar pairs of electrodes. This summation can exacerbate interactions among channels, especially for patients

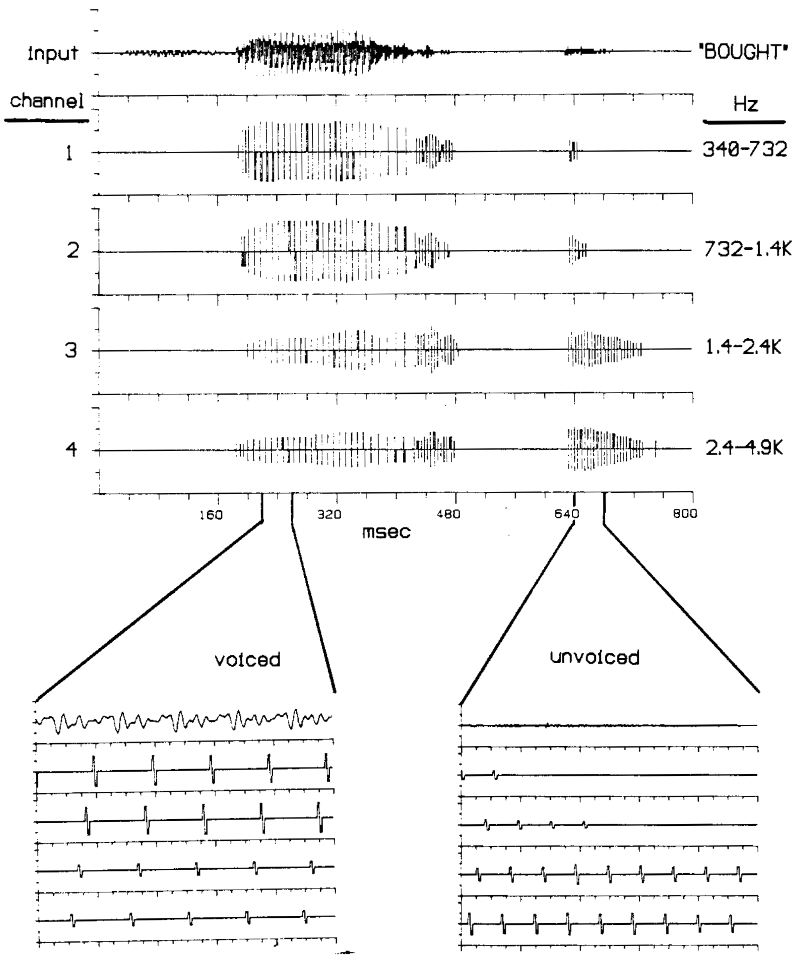


Fig. 2. Waveforms of an interleaved pulses (IP) processor.

who require high stimulation levels. Summation of stimuli from multiple channels also depends on the phase relationships among the waveforms. Because these relationships are not controlled in a multichannel CA processor, representation of the speech spectrum may be further distorted by continuously changing patterns of channel interaction. A reduction of channel interactions might increase the salience of channel-related cues for implant patients.

The problem of channel interactions is addressed in the IP processor of Fig. 2 through the use of nonsimultaneous stimuli. There is no temporal overlap between stimulus pulses, so that direct summation of electrical fields produced by different electrode channels is avoided. The energy in each frequency band of the input signal is coded as the amplitude of the pulses delivered to the corresponding stimulus channel. Distinctions between voiced and unvoiced segments of speech are represented by the timing of cycles of stimulation across the electrode array. In this particular processor stimulation cycles are timed to occur in synchrony with the detected fundamental frequency for voiced speech sounds and at the maximum rate (with one stimulation cycle immediately following its predecessor) for unvoiced speech sounds. The timing of stimulation cycles for voiced and unvoiced intervals can be seen in the lower panels of Fig. 2.

Comparison of Figs. 1 and 2 shows large differences in the stimuli presented for the CA and IP processing strategies. One might expect that the CA processing strategy would provide the greatest benefits to patients who could appreciate details in the stimulus waveforms (see above) and who have low thresholds of stimulation. In contrast, the IP processing strategy might be expected to provide superior performance for less fortunate patients who cannot make use of such details in CA stimulus waveforms and who have high thresholds of stimulation. We note that high thresholds of stimulation and high levels of measured channel interactions with simultaneous stimuli are both regarded as signs of poor nerve survival in the implanted ear [Gardi, 1985; Merzenich et al., 1978; Pfingst and Sutton, 1983; White et al., 1984]. Thus, application of an IP processor may confer special benefits for patients with poor nerve survival.

Subjects and Tests

All eight patients participating in these studies had been implanted with the UCSF electrode array. The first two patients were fitted with percutaneous cables and the remaining six with the four-channel transcutaneous transmission system (TTS) of the UCSF/Storz prosthesis. Use of the cable allows direct access to all 16 electrode contacts in the array (usually configured as eight bipolar pairs) and direct control over the current or voltage waveforms of the stimuli. In contrast, alternating bipolar electrode pairs are assigned to the four channels of the TTS, and the current and voltage waveforms are complex functions of the nonlinear impedances of the electrodes and the limited frequency response of each TTS channel.

Because the cable allowed a much greater degree of stimulus control, many different processing strategies were evaluated in the studies with the first two patients (LP and MH). The performance of each strategy was evaluated with tests of consonant and vowel identification. The vowel test included the tokens "BOAT," "BEET," "BOUGHT," "BIT," and "BOOT," and the consonant test included the nonsense tokens "ATA," "ADA," "AKA," "ASA," "AZA," "ANA," "ALA," and "ATHA." The vowel tokens were selected to measure the ability to perceive differences in the formant frequencies of these particular vowels, and the consonant tokens were selected to measure the ability to distinguish the nonlabial consonants that have the greatest frequency of occurrence in spoken English [Schubert, 1985].

Among the processing strategies evaluated with these first two patients, the IP approach achieved the best results. To assess further the performance of this strategy *vis à vis* the standard CA strategy of the UCSF/Storz prosthesis, an extensive series of standard tests was designed for the remaining six patients (all of whom were fitted with the TTS). In addition to the vowel and consonant confusion tests just outlined, this series included:

all subtests of the Minimal Auditory Capabilities (MAC) battery [Owens et al., 1985]; the Diagnostic Discrimination Test (DDT) of consonant confusions [Grether, 1970]; connected discourse tracking with and without the prosthesis [De Filippo and Scott, 1978; Owens and Raggio, 1987]; and the Iowa test of medial consonant identification with visual cues [Tyler et al., 1983]. In this report we will review briefly the results from the vowel and consonant confusion tests for all patients, and from the MAC and Iowa tests for the six patients fitted with the TTS.

Results for Patient LP

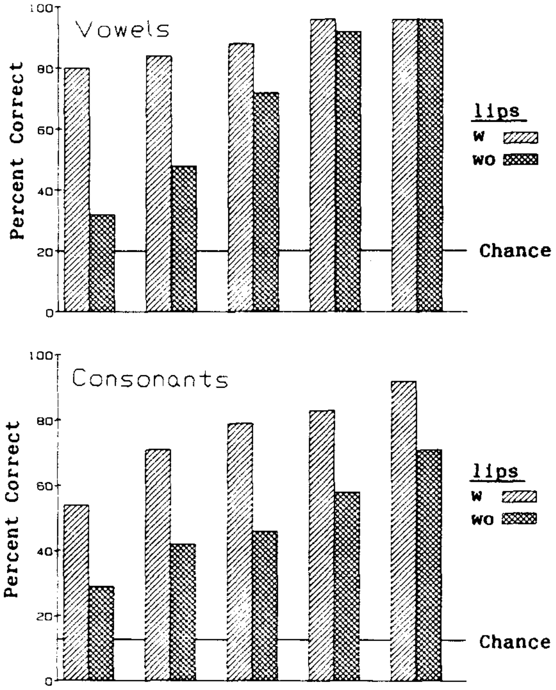
The first patient (LP) had a most discouraging picture of psychophysical performance. He had extremely severe channel interactions and high thresholds for bipolar electrical stimulation. His case was further complicated by extraordinarily narrow dynamic ranges and lability of thresholds and loudness levels both within and between testing sessions. LP's psychophysical findings of severe channel interactions, high thresholds and narrow dynamic ranges were all consistent with a picture of very poor survival of neurons in his implanted ear [Gardi, 1985; Merzenich et al., 1978; Pfingst et al., 1985; Pfingst and Sutton, 1983; White et al., 1984; Wilson et al., 1988a].

As might be expected, LP received no benefit from the CA processor used in the standard UCSF prosthesis. Indeed, he refused to describe any of the percepts produced with this processor as speechlike.

The first application of a 6-channel, IP processor immediately moved LP into the speech mode of auditory perception [Wilson et al., 1985; 1988a]. Of the eleven vowel and consonant tokens initially presented to LP using the IP processor, seven were spontaneously recognized as the correct words or syllables. Although his performance declined when the number of channels was reduced from 6 to 4, formal tests of vowel identification indicated that LP could perform at a level significantly above chance ($p < 0.01$) even with a reduced 4-channel version of the IP processor [Wilson et al., 1985]. A medical complication required surgical removal of LP's implant, ending our study of him with these very encouraging preliminary results.

Results for Patient MH

With the second patient we were able to evaluate differences in processor performance in much greater detail. This patient also had psychophysical manifestations of poor nerve survival [Wilson et al., 1986a, 1988a]. The results presented in Fig. 3 show her levels of performance in tests of vowel and consonant identification. The diagonally hatched bars show her performance with lipreading, and the cross hatched bars without. The chance levels of performance are indicated by the horizontal lines in each panel. Different processors are represented by different sets of bars. The characteristics of each processor are indicated in the labels at the bottom. For example, the leftmost set of bars shows the scores for a 4-channel CA processor. The remaining sets of bars show the scores for four variations of IP processors. These variations were produced by manipulating (a) the number of stimulation channels and (b) the way in which the beginnings of stimulus sequences were timed. In one approach stimulation cycles were timed to start



Proc:	CA	IP	IP	IP	IP
Chans:	4	4	4	6	6
FO, v/uv?	NA	N	Y	N	Y
Overall %	49	61	71	83	89

Fig. 3. Results of vowel and consonant confusion tests for patient MH. Diagonally-hatched bars indicate results obtained with lipreading and cross-hatched bars indicate results obtained without lipreading. The table at the bottom of the figure indicates the type of processor used (abbreviations are CA for "compressed analog" and IP for "interleaved pulses"); the number of stimulation channels; whether voicing information was explicitly coded for the IP processors; and the overall percent-correct scores from the four test conditions for each processor. The horizontal line in each panel shows the level of chance performance for that test. All of the scores are significantly above chance ($p < 0.01$), except those for the CA processor tests without lipreading.

in synchrony with the fundamental frequency for voiced speech sounds and at randomly-spaced intervals during unvoiced speech sounds. This constituted explicit coding of fundamental frequency and voiced/unvoiced distinctions. In the other approach stimulation cycles were timed to follow each other as rapidly as possible, providing no explicit coding of voicing information. In all, the results shown in Fig. 3 allow direct comparisons of (a) a 4-channel CA processor *versus* a 4-channel IP processor; (b) 4- *versus* 6-channel IP processors; and (c) IP processors with and without explicit coding of voicing information. The comparisons indicate that:

1. Performance is markedly improved when a 4-channel IP processor is used instead of a 4-channel CA processor;

2. Scores are *much* higher in all categories except vowel identification with lipreading (where scores are about the same) when a 6-channel IP processor is used instead of a 4-channel IP processor; and

3. Explicit coding of voicing information improves the performance of IP processors, particularly in the categories of vowel identification without lipreading (4-channel processor), consonant identification without lipreading (6-channel processor) and consonant identification with lipreading (both processors).

Good test/retest reliability was thoroughly demonstrated for MH. When retested with a processor that had produced low scores on a previous occasion, MH always would obtain low scores again, and when repeating a test with a processor that earlier had performed well she always would repeat her high scores. The standard deviation of overall percent correct scores from seven repeated trials of the last (rightmost) processor shown in Fig. 3, for example, was slightly less than three percent. All scores presented in Fig. 3 are significantly above chance ($p < 0.01$) except for those of the two hearing-alone conditions for the CA processor.

Results for Patients Fitted with the TTS

The studies with the two percutaneous cable patients demonstrated that (a) different processing strategies can produce widely different outcomes for individual patients and (b) IP processors are far superior to the tested alternative processors for at least two patients with psychophysical signs of poor nerve survival. With these observations in mind, we were most interested in comparing the IP and CA strategies in extensive tests with a larger population of patients. We wondered, for example, how the IP processor would perform for successful users of a CA processor, and whether the potential advantages of the IP processors could be realized in patients with four or fewer channels of stimulation.

Six patients implanted with the UCSF/Storz cochlear prosthesis [Loeb et al., 1983; Merzenich, 1985; Schindler et al., 1986; Wilson et al., 1988d] participated as subjects in the follow-up study. Tests with the CA processor were conducted with each patient's clinical device, and tests with the IP processor were conducted either with computer simulations [Wilson and Finley, 1985] or with a real-time, microprocessor-based instrument [Finley et al., 1987].

Each subject was studied for a one-week period in which (a) basic psychophysical measures were obtained on thresholds and dynamic ranges for pulsatile stimuli, (b) a variety of IP processors (with different choices of processor parameters) was evaluated with tests of vowel and consonant identification, and (c) the best of these IP processors was further evaluated using a broad spectrum of speech tests.

It is important to note that certain attributes of these subjects favored the CA processor in comparisons of the CA and IP strategies. First, all six subjects entered the study with substantial experience using the CA processor. The average experience with this processor approximated 1 year of daily use. In contrast, experience with the IP processor was limited to that obtained in a 6-day period of testing a variety of processors with each subject. As mentioned before (and discussed in detail elsewhere -- see Dowell et al., 1987; Tyler et al., 1986; Wilson et al., 1988d), such a disparity in experience might strongly favor the CA processor.

An additional factor weighing against the IP processor was the use of UCSF/Storz TTS. The principal limitations of that system for IP processors were (a) inadequate levels of voltage compliance for stimulation with short-duration pulses, (b) the small number of channels, and (c) lack of current control in the stimulus waveforms. Half of the subjects were further limited to fewer than four channels due to a mode of device failure [Schindler et al., 1986]. Because optimized fittings of IP processors require at least six channels of stimulation and short-duration pulses [Wilson et al., 1988a,b], compromises had to be made in the fitting of IP processors for the subjects of this study.

The parameters selected for the IP processors used by each of the six subjects are presented in Table 3. The best fulfillments of the fitting criteria for IP processors [Wilson et al., 1988b] were obtained for subjects HE and MC2. Each had the use of all four stimulation channels, and the average pulse width across channels was 0.5 ms/phase or less for these two subjects.

In contrast, relatively poor sets of parameters had to be used for the remaining subjects. Subjects MC1 and JM had only three usable channels and subject RC only two. In addition, long pulse durations (1.0 ms) had to be used for subjects JM and ET.

The results of processor comparisons are presented in Figs. 4-8. Fig. 4 shows performance levels for the tests of vowel and consonant identifications. Solid black bars represent the scores for lipreading alone, dotted bars for the CA processor alone, and cross-hatched bars for the CA processor with lipreading. Vertically-lined bars represent the scores for the IP processor alone and horizontally-lined bars for the IP processor with lipreading. Chance levels of performance are indicated by the horizontal lines in each panel and, as is the case in subsequent figures, any results that do not significantly exceed chance are noted in the figure captions.

Performance in the tests of vowel identification is quite high for both processors. All patients have scores of 80% correct or better for both processors with lipreading and 68% or better without lipreading. The means of the results for the two processors are not significantly different for either condition.

In contrast to the apparent equivalence of processors for the vowel test, the scores for the consonant test without lipreading are significantly higher for the IP processor ($p < 0.05$, paired t test). All patients except JM

Table 3. Parameters of IP processors.*

Subject	Channels	Pulse Widths/Phase (ms)	Pulse Sep. (ms)	Cycle Time (ms)
MC1	3	0.5	0.5	4.5
HE	4	0.5	0.5	6.0
JM	3	1.0	0.1	6.3
RC	2	0.5	0.1	2.2
ET	4	1.0 1.0 0.5 0.5	0.1	6.4
MC2	4	0.3 0.7 0.3 0.3	0.5	5.2

*All six processors used symmetric biphasic pulses with positive phase leading and with the channels stimulated in base-to-apex order. Stimulation cycles were presented at the fundamental frequency for voiced intervals and at maximum rate (period equal to cycle-time) during unvoiced intervals.

obtained higher scores using the IP processor for this condition. The difference between the means for the "with lipreading" condition is not significant. Analyses of confusion patterns with each processing strategy are presented elsewhere [Wilson et al., in press].

An indication of superior consonant recognition with the IP processor also is seen in the results for the Iowa test of medial consonant identification with lipreading cues. These results are presented in Fig. 5. Every subject except MC2 has substantially higher scores with the IP processor. However, the difference in the means of the results for the two processors is not statistically significant. It is important to note that our procedure for administering this test was designed to confer any benefits of learning on the CA processor. The order of testing was to measure performance first with the IP processor plus lipreading, then with lipreading alone, and finally with the CA processor plus lipreading.

In addition to paired t tests, a Wilcoxon test (see Hollander and Wolfe, 1973) for significant improvement across the six patients was applied to each subtest of our combined battery. Only two tests emerged from the Wilcoxon

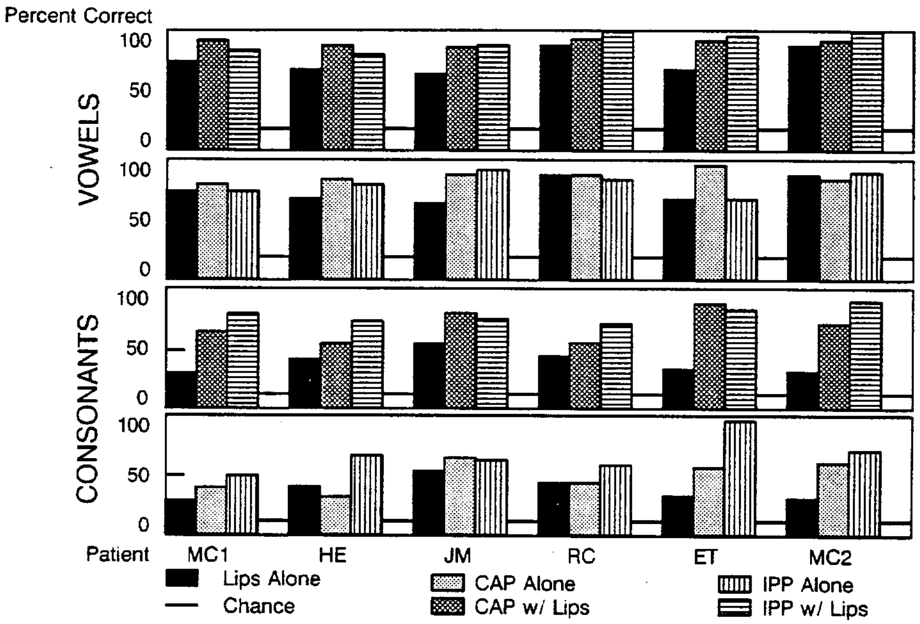


Fig. 4. Results from the RTI tests of vowel and consonant identification for six subjects implanted with the UCSF/Storz cochlear prosthesis. All scores are significantly above chance ($p < 0.01$), except for some scores for the consonant tests with lipreading alone.

analysis as indicating significant improvement with the use of one processor over the other. Both the RTI consonant confusion test without lipreading and the Iowa test of medial consonant identification with lipreading indicated better performance with the IP processor (Wilcoxon T+ = 20, p = 0.03).

The remaining results presented in Figs. 6-8 are those from the MAC battery. Results from the subtests of prosodic perception (timing of syllable boundaries, voice fundamental frequency and word stress) are shown in Fig. 6. These results demonstrate that, in general, these subtests are too easy for the patients and processors under study. The scores for the noise/voice (N/V) and spondee same/different (Sp S/D) tests are high for all patients and for both processors. The accent test is a more sensitive indicator of performance. Scores for this test show clear differences among patients, but relatively modest differences between processors. Three patients did moderately better with the IP processor on this test (RC, ET, and MC2), and two patients did moderately better with the CA processor (MC1 and JM). Finally, for the Question/Statement (Q/S) test one patient did much better with the CA processor (MC1), three patients did somewhat better with the IP processor (HE, JM and RC) and two patients obtained identical scores with the two processors (ET and MC2). None of the differences in the means of the results for the two processors is statistically significant among the prosodic subtests of the MAC battery.

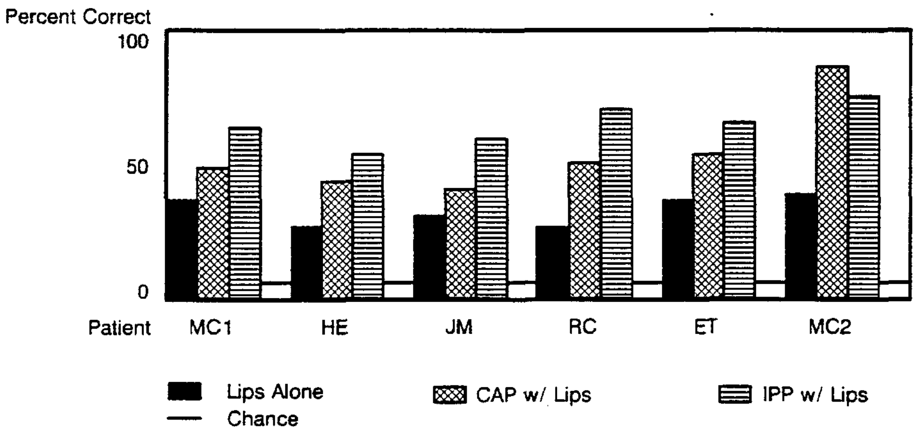


Fig. 5. Results from the Iowa test of consonant identification with lipreading cues. The subjects are those described in the caption of Fig. 4. All scores are significantly above chance ($p < 0.01$).

Results from the phoneme and word subtests of the MAC battery also demonstrate a general equivalence of the processors for the conditions of our study. These results are presented in Fig. 7. Again, none of the differences in the means of the results for the two processors is statistically significant. Patients HE and MC2 have somewhat higher or equivalent scores on all four tests (including vowel, initial consonant, final consonant, and four-choice spondee) with the IP processor, and patients MC1 and RC have somewhat higher or equivalent scores with the CA processor.

Finally, results from the subtests of open-set recognition are shown in Fig. 8. Once again, none of the differences in the means of the results for the two processors is statistically significant. Patient MC2, however, has higher scores for the IP processor for all four subtests of spondee recognition (Sp), recognition of monosyllabic words (NU6), recognition of CID sentences (CID), and recognition of words in context (WIC). In addition, patient HE has much higher scores with the IP processor for the tests of spondee recognition and of words in context. On the other hand, patient RC has higher scores with the CA processor on every open-set test, and patients MC1 and ET show generally superior performance with the CA processor for those tests.

Summary of Results

In the studies with the two percutaneous cable patients, each of whom had a psychophysical picture consistent with poor nerve survival, the IP

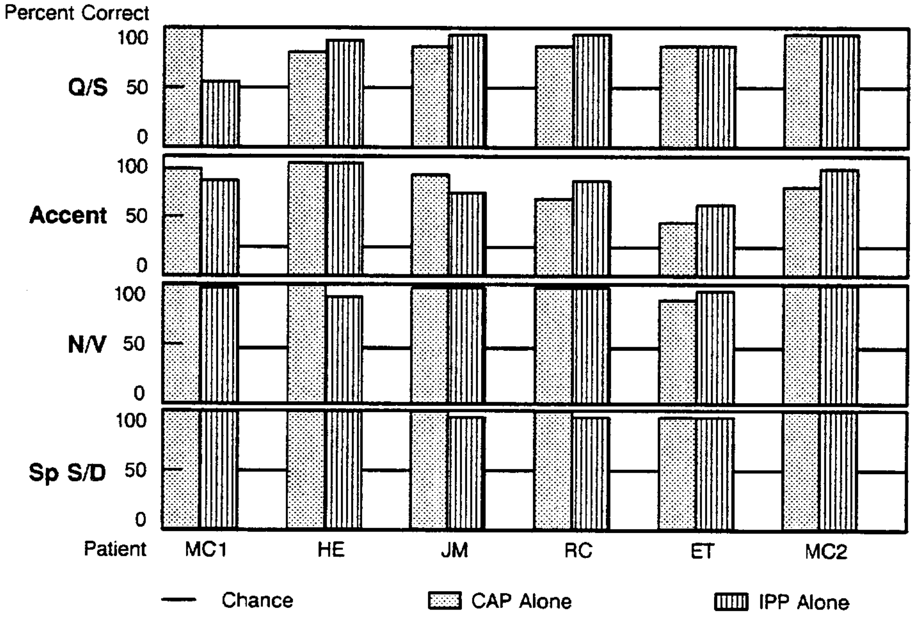


Fig. 6. Results from subtests of the MAC battery designed to evaluate perception of prosodic elements in speech. Abbreviations for the subtests are Q/S for question/statement; N/V for noise/voice; and Sp S/D for spondee same/different. The subjects are those described in the caption of Fig. 4. All scores are significantly above chance ($p < 0.01$), except MC1's score for the Q/S test with the IP processor and ET's score for the accent test with the CA processor.

strategy produced much better results than the CA strategy. We believe this improved performance with the IP processor is attributable in part to the sizable reduction in channel interactions afforded by the use of nonsimultaneous stimuli. The studies with the cable patients further demonstrated a strong correspondence between number of stimulation channels and performance. In the studies with the six patients fitted with the TTS, the IP and CA processors were compared under conditions of substantial experience with the CA processor, generally high levels of performance with the CA processor, and severe restrictions imposed by the TTS for implementing optimized versions of the IP processor. Despite these limitations, two patients in this second series *immediately* had better performance with the IP processor (patients HE and MC2) and three had similar or slightly inferior levels of performance with this processor (patients MC1, JM and ET). Only one patient (RC) had clearly superior performance with the CA processor. This patient also happened to have the highest level of performance among the six studied patients with the CA processor and, with only two functioning channels, afforded the poorest fulfillment of our IP processor fitting criteria.

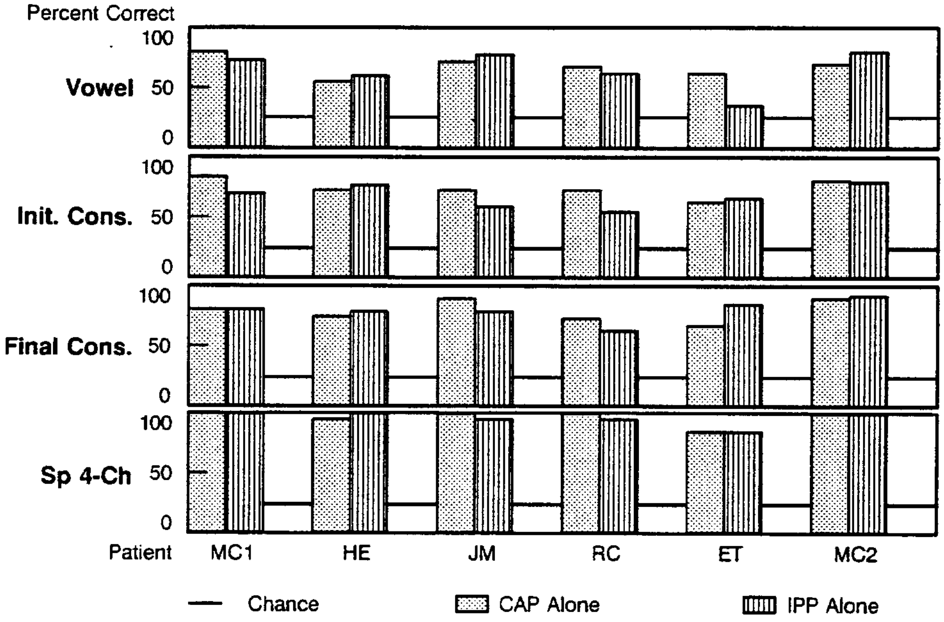


Fig. 7. Results from subtests of the MAC battery designed to evaluate perception of phonemes and words in speech. Abbreviations for the subtests are Init. Cons. for initial consonant; Final Cons. for final consonant; and Sp 4-Ch for four-choice spondee. The subjects are those described in the caption of Fig. 4. All scores are significantly above chance ($p < 0.01$), except ET's score for the vowel test with the IP processor.

These results suggest that (a) most patients are likely to obtain at least equivalent results if an IP processor is used instead of a CA processor; (b) patients with psychophysical signs of poor nerve survival are likely to obtain better results with an IP processor; and (c) use of a TTS designed to support an IP processor (e.g., a TTS with eight channels of current-controlled outputs) is likely to produce results that are better than those obtained with the limited TTS of the present studies.

Discussion

A key finding of the studies reviewed above is that substantial gains in speech perception can be made by selecting an appropriate processing strategy for each patient. In our two series of patients RC obtained clearly superior results with the CA processing strategy while LP and MH obtained clearly superior results with the IP processing strategy. These results demonstrate that access to a variety of alternative strategies may be required for optimizing outcomes across a population of patients.

The subject who obtained superior results with the CA processor (RC) had

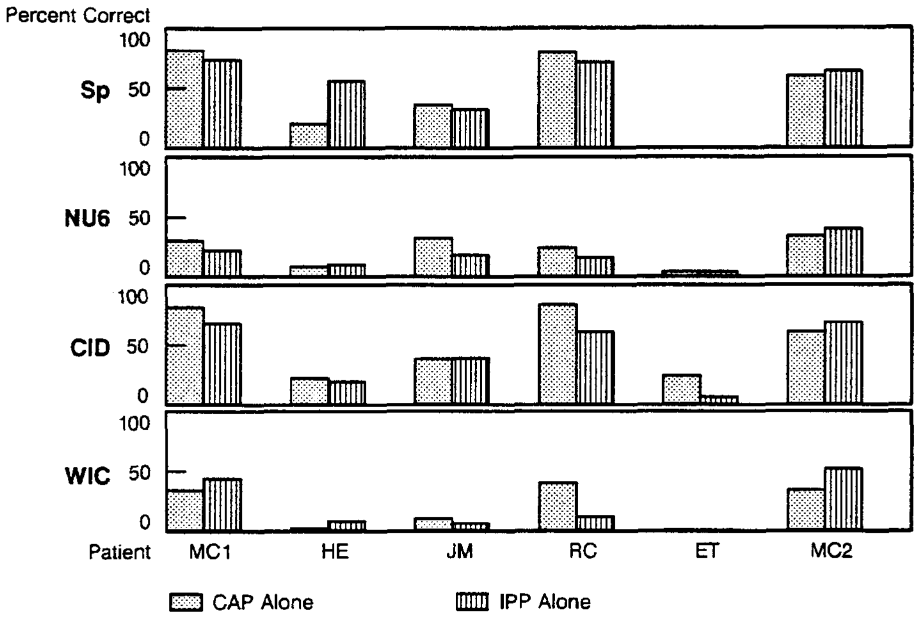


Fig. 8. Results from subtests of the MAC battery designed to measure open set recognition of speech. Abbreviations for the subtests are Sp for spondee recognition; NU6 for recognition of monosyllabic words from Northwestern University list 6; CID for recognition of everyday sentences from lists prepared at the Central Institute for the Deaf; and WIC for recognition of words in context. The subjects are those described in the caption of Fig. 4.

only two functional channels of intracochlear stimulation. This number is certainly too few for even a gross representation of the speech spectrum with an IP processor. The relatively poor performance of the IP processor therefore could be attributed at least in part to a poor match to its fitting criteria.

Another possible explanation for RC's superior performance with the CA processor is that he made especially good use of the information present in the CA stimulus waveforms. Indeed, the impressive results obtained with RC (two channels), MC1 (three channels), certain patients in the Vienna series (one channel; see Hochmair-Desoyer and Burian, 1985; Hochmair-Desoyer et al., 1985), and certain patients in the Symbion series (four monopolar channels with relatively poor isolation; see Eddington, 1983) support the hypothesis that the major bearer of information in CA processors is the waveform itself. Although results from studies conducted at UCSF demonstrate that additional information can be provided with four channels of CA stimulation using the UCSF electrode array [Ochs et al., 1985; Schindler et al., 1987; White et al., 1985], this additional information is not required for excellent performance in some patients. Most likely, the best results are obtained for patients who have the greatest access to information in the CA waveform(s). These patients might include those with exceptional abilities to discriminate (a) frequencies up through the range of the first formant of speech [Eddington, 1983; Hochmair-Desoyer and Burian, 1985; White, 1983]; (b) rapid temporal variations in the envelopes of speech and speech-like stimuli [Hochmair-Desoyer et al., 1985; Soli et al., 1986; Van Tasell et al., 1987]; and (c) subtle waveshape changes produced by the addition of frequency components beyond the first formant [Dobie and Dillier, 1985; Hochmair and Hochmair-Desoyer, 1985].

If this second interpretation is pertinent to RC's case, then patients with such special abilities might be best served with a CA processor. Optimal implementations of this processor would provide any additional information the patient might be able to utilize in multiple channels of stimulation. The maximum number of useful channels is likely to be limited, however, by the severe interactions that can occur among channels when simultaneous stimuli are used.

In contrast, patients with a psychophysical picture consistent with poor nerve survival (i.e., severe channel interactions, high thresholds, narrow dynamic ranges, and perhaps limited abilities to discriminate frequencies and other stimulus attributes) are likely to receive greater benefit from an IP processor. For two such patients in our initial series, large increases in performance were obtained when the number of stimulus channels was increased from 2 to 4 and from 4 to 6. Paradoxically, then, multichannel implants may provide relatively greater benefits to patients with signs of poor nerve survival than to patients without these signs.

To summarize the comparisons made above, Table 4 lists characteristics of the CA and IP processing strategies. Briefly, the CA strategy may be superior for patients with good nerve survival because such patients may perceive substantial temporal and frequency information in analog waveforms and because the lower stimulus intensities required for these patients, along with survival of neural elements over the active electrodes, can greatly minimize channel interactions produced by simultaneous stimulation. On the other hand, the IP strategy may be superior for patients with poor nerve survival because isolation between channels for such patients is improved with the use of nonsimultaneous stimuli.

Finally, we note that these comparisons between processors suggest possibilities for further improvements in performance. One such improvement might be made by combining the best features of the CA and IP approaches in "hybrid" strategies. For a good nerve survival case, for example, the main benefits of the CA strategy might be realized with a single channel of stimulation. This would leave the remaining channels for the representation of frequency components in speech above the first formant. The excellent results obtained in the present studies with all eight patients using the IP processor (especially patients HE, MC2 and MC1) indicate that interleaved stimuli are likely to enhance speech representation even for patients with good nerve survival (presumably through further reduction of interactions between adjacent channels). The combined use of the CA and IP strategies therefore might confer in an optimal way the benefits of waveform discrimination and multichannel stimulation to fortunate patients with good

Table 4. Characteristics of Processors.*

ANALOG	PULSATILE
continuous waveforms, presented simultaneously	nonsimultaneous pulses
severe interactions between channels for patients with poor nerve survival	improved channel isolation, especially for patients with poor nerve survival
in some patients, continuous waveforms can provide good temporal and frequency information (F0, voiced/unvoiced distinctions, F1, possibly F2)	limited transmission of temporal and frequency information (F0, voiced/unvoiced distinctions)

*Symbols used in this Table are F0 for the fundamental frequency of voiced speech, F1 for the first formant frequency of speech, and F2 for the second formant frequency of speech.

nerve survival. Similarly, for cases in which nerve survival is patchy, psychophysical or electrophysiological tests might be conducted to identify areas of good survival. A bipolar pair of electrodes adjacent to one of these areas could be selected for compressed analog stimulation. This electrode channel would have low threshold and suprathreshold stimulus levels relative to electrodes adjacent to poor survival areas. Such low levels might allow the remaining electrode channels to receive IP stimuli with only minor channel interactions.

Potential applications of hybrid processors, along with the choices posed by the existing CA and IP strategies, emphasize the need for flexibility in the fitting of speech processors to individual patients. We believe further significant advances in the development of speech processors for cochlear prostheses will result from (a) an improved understanding of the electrode-nerve interface, especially as it relates to the pattern of nerve survival, and (b) design and use of better psychophysical and electrophysiological tests to infer the pattern of survival in the implanted ear.

Conclusions

Major conclusions from the studies reviewed above include the following:

1. Different processing strategies can produce widely different outcomes for individual implant patients;

2. IP processors are far superior to the tested alternative processors for at least two patients with psychophysical signs of poor nerve survival;

3. The performance of IP processors strongly depends on the selection of processor parameters;

4. Use of a TTS designed to support an IP processor (e.g., a TTS with eight channels of current-controlled outputs) is likely to produce results that are better than those obtained with the limited TTS of the present studies;

5. Processors other than the IP processors can be superior for patients with psychophysical signs of good nerve survival and for patients who cannot be fitted with an optimized IP processor;

6. One such processor is the CA processor of the present UCSF/Storz cochlear prosthesis; and

7. Substantial gains in speech understanding can be made by (a) selecting the best type of speech processor for each patient and (b) using implanted and external hardware capable of supporting a wide range of different processing strategies.

III. Design and Evaluation of a Two-Channel, "Breeuwer/Plomp" Processor

In a recent paper Breeuwer and Plomp described a speech processor which was particularly effective as an aid to lipreading [Breeuwer and Plomp, 1984]. The supplementary signal provided by this processor consisted simply of acoustic representations of the root-mean-square (RMS) energies in a pair of octave bands centered at 500 and 3160 Hz. To evaluate the processing strategy, Breeuwer and Plomp measured the number of correctly perceived syllables in short Dutch sentences presented to 18 listeners with normal hearing. The test conditions included lipreading only, lipreading plus acoustic supplement, and acoustic supplement alone. The results were impressive. The percentage of correctly perceived syllables jumped from 23% correct for lipreading only to 87% correct for lipreading plus the processed speech supplement, a score fully consistent with substantial open-set recognition of speech.

These results led us to evaluate a modified "Breeuwer/Plomp" (B/P) strategy in tests with cochlear implant patients [Wilson et al., 1987]. One motivation for this study was to determine the efficacy of the B/P strategy in situations allowing the use of only a few channels of electrical stimulation. Such situations are surprisingly numerous and include (a) the use of electrode placements or configurations with inherently poor isolation between channels, as in extracochlear auditory prostheses; (b) the use of two-electrode devices, as is presently the case for stimulation of the cochlear nucleus [Eisenberg et al., 1987]; and (c) cases in which only a few channels of stimulation in a multichannel intracochlear device can be perceived independently due to poor survival of cochlear neurons, device failure (see Wilson et al., 1988d), or partial insertion of the electrode array.

Subjects

Five cochlear implant patients participated as subjects. They included MC1, HE, RC, ET and MC2 from the studies described in section II of this report. As indicated in that section, each of these patients was implanted with the UCSF/Storz electrode array and fitted with the four-channel UCSF/Storz system for transcutaneous transmission of stimulus information across the skin. Patient JM was unable to participate in the evaluations of the B/P processor due to lack of time.

Processing Strategy

Our application of the processing strategy described by Breeuwer and Plomp consisted of mapping the RMS energies of the two octave-wide bands into the dynamic range of electrically evoked hearing for two channels of intracochlear stimulation. The stimuli were interleaved pulses derived in a manner identical to that described in section II. The only differences between the two-channel B/P processors of the present section and the three- or four-channel IP processors of section II were (a) the use of two octave-wide bands for the B/P processor, as opposed to contiguous bands spaced along a logarithmic scale for the IP processors, and (b) the lower number of stimulation channels for the B/P processor.

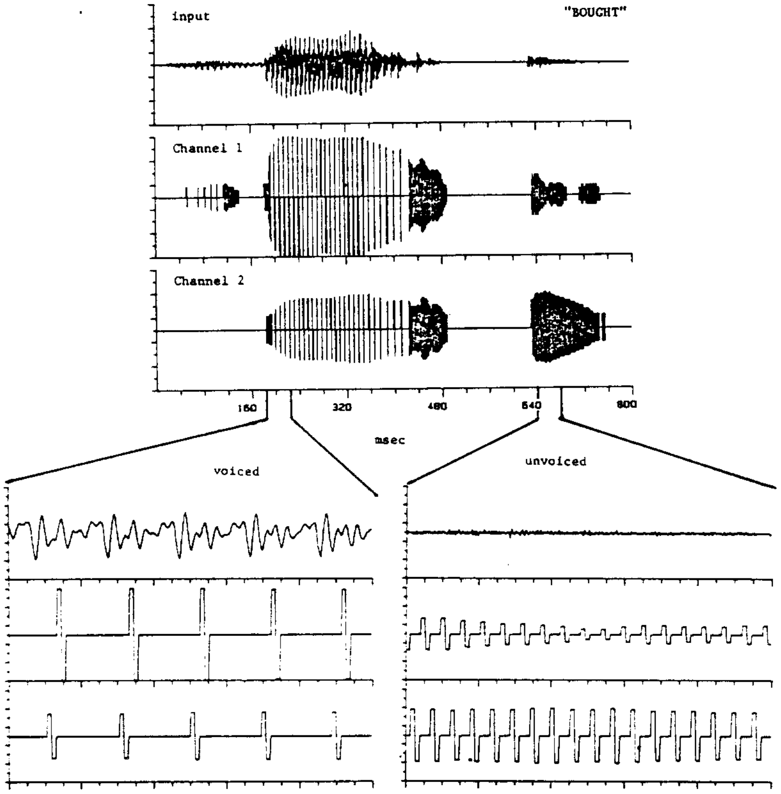


Fig. 9. Waveforms of a two channel, "Breeuwer/Plomp" (B/P) processor.

Typical waveforms for our implementation of the B/P processor are shown in Fig. 9. The top trace in each panel is the processor input and the remaining traces are channel outputs. The input is the word "bought." The initial consonant occurs at about 180 ms and the vowel follows immediately thereafter. An expanded display of waveforms well into the vowel is shown in the lower left panel. The /t/ burst begins slightly before 640 ms, and an expanded display of waveforms beginning at 640 ms is shown in the lower right panel.

In the particular variation of B/P processors presented in Fig. 9, balanced biphasic pulses are used, and voicing information is explicitly coded. During voiced speech the stimulation cycles are timed to begin in synchrony with the detected fundamental frequency, while during unvoiced speech the stimulation cycles are initiated as rapidly as possible (with one stimulation cycle immediately following its predecessor). The pulse sequence in each stimulation cycle is such that the more basal electrode channel is stimulated first.

Methods

The parameters selected for the B/P processors used in this study are presented in Table 5. The electrode channels chosen for each subject provided

Table 5. Parameters selected for Breeuwer/Plomp processors.

subject	channels[a]	pulse duration/ph[b]	pulse sep.[b]	cycle time[b]	Max Rate or Jittered
MC1	1,5	0.5	0.5	3.0	MR
HE	1,7	0.5	0.5	3.0	MR
RC	5,7	0.5	0.1	2.2	MR
ET	1,7	0.5	0.5	3.0	MR
MC2	1,5	0.3	0.5	2.2	J

[a] Channels are numbered in ascending order from the apical end of the electrode array. Channel 1 corresponds to bipolar electrode pair 1+2, and channel 8 to bipolar electrode pair 15+16.

[b] Times are in milliseconds.

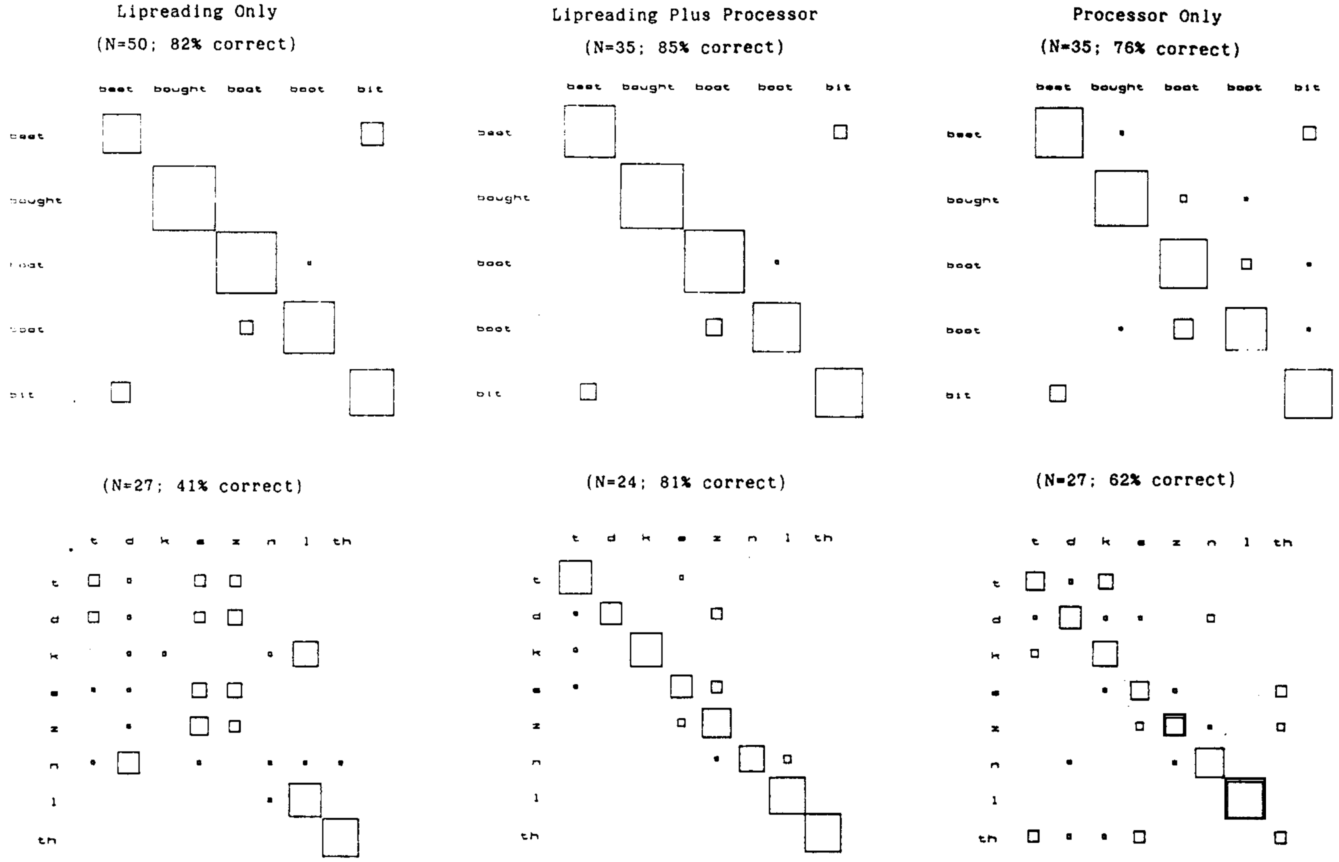
distinct "place pitch" percepts (i.e., differences in timbre or pitch for loudness-balanced stimuli allowed reliable discrimination of the selected channels for all subjects). In addition, the processors for all subjects used explicit coding of voicing information. Stimulation cycles were initiated at the maximum rate during unvoiced intervals for subjects MC1, HE, RC and ET, and stimulation cycles were initiated at randomly-spaced intervals for subject MC2.

Tests of vowel and consonant identification were conducted for all five subjects. These tests were identical to those described above for the comparisons of the CA and IP processing strategies (see section II).

In addition to the vowel and consonant tests, the B/P strategy was further evaluated with an extensive battery of speech perception tests for one of the subjects. These tests included all subtests of the Minimal Auditory Capabilities (MAC) battery [Owens et al., 1985], connected discourse tracking with and without the processor [De Filippo and Scott, 1978; Owens and Raggio, 1987], and the Iowa test of medial consonant identification with video presentations of the speaker's face [Tyler et al., 1983].

Results

Results from the tests of vowel and consonant identifications are presented in Fig. 10 in the form of confusion matrices. The responses from all five subjects are combined in the matrices for each condition. A significant improvement in consonant identification is produced when the auditory cues from the B/P processor are added to the visual cues provided by lipreading ($p < 0.002$, paired t test). Especially impressive is the improvement in the identification of the consonants that are least visible on the lips. Examination of the matrix for consonant identification with


Lipreading Only
(N=50; 82% correct)

Lipreading Plus Processor
(N=35; 85% correct)

Processor Only
(N=35; 76% correct)

(N=27; 41% correct)

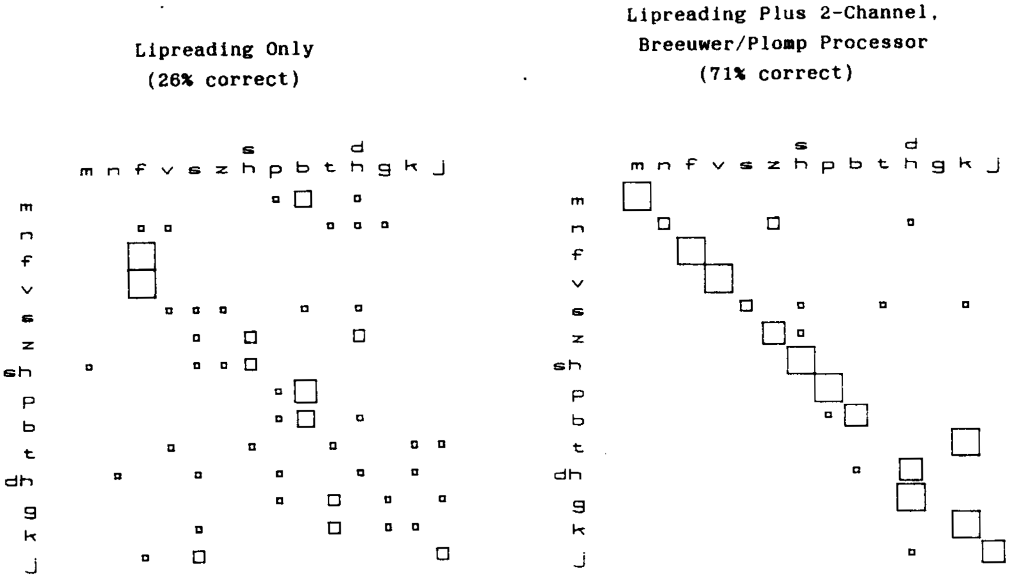
(N=24; 81% correct)

(N=27; 62% correct)

Fig. 10. Results from the RTI tests of vowel and consonant identification. The lengths of the sides of each square represent the averages of the scores obtained for the indicated conditions across all five subjects. Rows in the matrices represent stimuli and columns the responses. The total number of stimulus presentations and the overall percentage of correct responses are shown in the parentheses above each matrix.

lipreading only shows that /l/ and /ð/ are highly visible on the lips while the other consonants are not. The percentage of correctly identified consonants other than /l/ and /ð/ is 22% for lipreading alone and 75% for lipreading plus processor. This improvement demonstrates the potential utility of the B/P processor as an adjunct to lipreading.

In addition, the B/P processor produced high vowel and consonant recognition scores with hearing alone. The overall scores were 76% correct for vowels and 62% correct for consonants. These scores are surprisingly good in view of Breeuwer and Plomp's report of 27% correct recognition of syllables for their hearing-only condition [Breeuwer and Plomp, 1984]. The better results in the present study might be attributable to the explicit coding of


Lipreading Only
(26% correct)

Lipreading Plus 2-Channel,
Breeuwer/Plomp Processor
(71% correct)

Fig. 11. Results from the Iowa test of consonant identification with lipreading cues. The matrix on the left shows the performance of patient RC for lipreading only, and the matrix on the right shows his performance for lipreading plus the auditory input provided by a Breeuwer/Plomp processor. Rows in the matrices represent stimuli and columns the responses. Each token was presented five times in each test.

voicing information in our implementations of a modified "B/P" processor, or to the relatively small numbers of tokens used in the vowel and consonant tests.

The results from the additional studies conducted with one of the subjects are presented in Fig. 11 and Table 6. Fig. 11 shows confusion matrices from the Iowa test of medial consonant identification with lipreading cues. These results are consistent with the previous results from the RTI consonant test. Patient RC obtained a score of 26% correct for lipreading alone on the Iowa test and a score of 71% correct for lipreading plus processor, an almost three-fold improvement in consonant identification. Moreover, the pattern of errors in the lipreading plus processor condition suggests that even better results could be obtained with training or learning. In particular, /t/ always was perceived as /k/ but never vice versa, and /g/ always was perceived as /d/ but never vice versa. These asymmetries in the errors indicate that information was available to make the distinctions, but was not being used by the subject.

Table 6. Results from the Minimal Auditory Capabilities (MAC) battery for subject RC.

Tests	Chance (%)	Score (%)[a]
Prosodic Perception (closed set)		
Question/Statement	50	95
Accent	25	80
Noise/Voice	50	98
Spondee Same/Different	50	95
Phoneme & Word Discrimination (closed set)		
Vowels	25.	62
Initial Consonants	25	55
Final Consonants	25	63
4-Choice Spondee	25	95
Open-Set Speech Recognition		
Spondees		72
Monosyllabic Words (NU 6)		16
Sentences (CID)		61
Words in Context (SPIN)		12

[a] All scores for the closed set tests (Prosodic Perception and Phoneme & Word Discrimination) are significantly above chance ($p < 0.01$).

Results from the tests of connected discourse tracking further confirm the findings of improved performance when the processor is used in conjunction with lipreading. RC's scores on the tracking tests were 4 words/minute for lipreading alone and 44 words/minute for lipreading plus processor. The additional auditory cues provided by the processor thus bring RC from a poor level of performance to a moderately good level of performance. In fact, his tracking rate with the processor is about half of the average rate obtained in tests with normal-hearing subjects [Owens and Raggio, 1987].

The remaining tests conducted with RC were those of the full MAC battery. The results, presented in Table 6, demonstrate substantial access to speech information with hearing alone. All scores from the closed set tests of prosodic information and of phoneme and word discrimination are significantly above chance ($p < 0.01$). Indeed, the scores on the Question/Statement, Noise/Voice, Spondee Same/Different and 4-Choice Spondee tests are all 95% correct or better. Surprisingly high scores also are obtained for the Accent (80% correct), Vowel (62% correct), and Final Consonant (63% correct) tests. These high scores are unexpected inasmuch as the B/P processor was specifically designed merely to present supplementary cues for lipreading.

Finally, the results from the tests of open-set recognition further indicate that the B/P processor can provide useful information even when the visual cues from lipreading are not available. In particular, high levels of performance are demonstrated in the tests of spondee (72% correct) and sentence (61% correct) recognition, and good levels of performance are demonstrated for the more difficult tasks of monosyllabic word recognition (16% correct) and recognition of single words in context (12% correct). These scores are comparable to the best results reported for open-set recognition using any type of auditory prosthesis (see, e.g., Gantz et al., 1988). We note, however, that scores on open-set tests are likely to reflect the characteristics of the subject as well as the quality of input provided by the speech processor. The present open-set scores therefore should be regarded as an indication of the potential of the B/P processor with hearing alone. That is, the information provided by this processor is adequate to support high levels of open-set recognition; however, all implant patients may not be able to utilize this information as effectively as RC.

Conclusions

The results of this study confirm and extend the findings of Breeuwer and Plomp. In particular, the present results from our evaluation of the "B/P" processors for cochlear implants are consistent with the following conclusions: (a) the B/P processor can act as a powerful adjunct to lipreading, as evidenced by large improvements in consonant identification for five subjects and in connected discourse tracking for a single subject; (b) this processor can produce fairly high levels of vowel and consonant identification with hearing alone; and (c) the performance with the one subject tested with the MAC battery further indicates that the B/P processor can provide useful information even when visual cues from lipreading are not available. These observations support the use of a B/P processor in cochlear prostheses whenever the number of channels available for distinct stimulation is restricted to two or three.

IV. Additional Processor Comparisons

In addition to the processor comparisons outlined in sections II and III above, studies were conducted in this project to

1. Analyze and compare the patterns of consonant and vowel identifications obtained with CA and IP processors (see below; Wilson et al., in press);
2. Conduct a preliminary evaluation of the hybrid CA/IP processor suggested in section II of this report (see Lawson et al., 1989);
3. Evaluate further the B/P processor in follow-up tests with UCSF/Storz patients and in tests with two patients implanted with the experimental multichannel prosthesis developed by 3M (see Soli, in press; Soli and Wilson, 1988);
4. Compare the IP processor with the multichannel, analog strategy of the experimental 3M device, in tests with a patient implanted with that device (see Soli, in press; Soli and Wilson, 1988);
5. Compare the IP processor with the CA processor of the Symbion device, in tests with a patient implanted with that device; and
6. Compare the IP processor with the feature-extraction strategy [Clark et al., 1987] of the Nucleus device, in tests with a patient implanted with that device.

Detailed information on studies relating to points 1-4 above may be found in the indicated references. Also, a summary of results from the studies of point 1 is presented in the remainder of this section. Results from the studies with the patients implanted with the Symbion and Nucleus devices are still being analyzed. These latter results will be presented in a future report when the analysis is completed.

Information Transmission Analysis

As indicated above, results from tests of consonant and vowel identification were analyzed to obtain further insights into the strengths and weaknesses of the CA and IP processing strategies. The subjects and tests were those described in section II of this report. Analyses of the identification patterns included the information transmission (IT) analysis of Miller and Nicely [1955] and the sequential information analysis (SINFA) of Wang and Bilger [1973]. In this report we will summarize results from the IT analysis. Detailed results from both analyses may be found in Wilson et al. [in press].

To evaluate the patterns of confusions (and correct responses) for the conditions of the study outlined in section II, the combined matrix for the responses of all subjects for each condition was used as an input to IT analysis. In this analysis the "relative transinformation" is calculated for selected articulatory or acoustic features of the phonemes in the

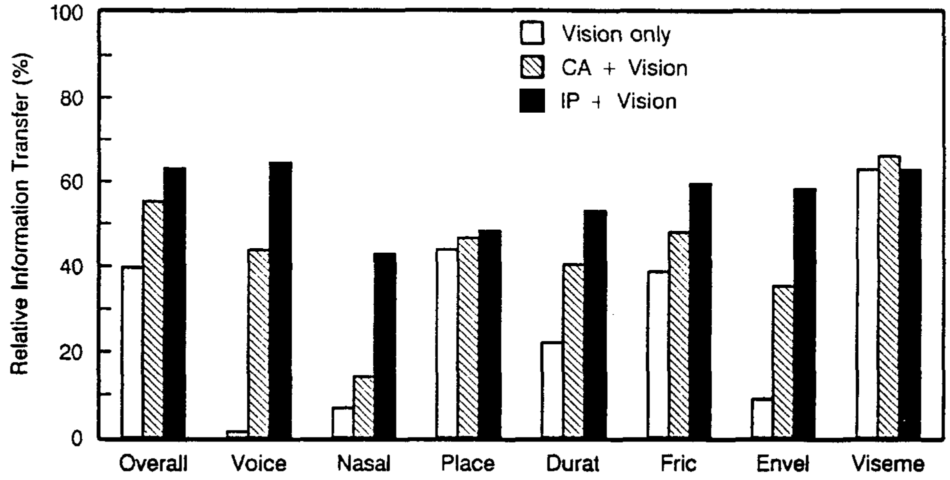


Fig. 12. Relative information transfer of speech features for the Iowa consonant test.

identification tests. The relative transinformation score for each feature, expressed here as percent information transfer, indicates how well that feature was transmitted to the subjects. The consonant features selected for the present study were voicing (voice), nasality (nasal), place of articulation (place), duration (durat), frication (fric), envelope cues (envel), and visual cues (viseme). The vowel features were first formant frequency (F1), second formant frequency (F2), duration (durat), and visual cues (viseme).

The results from IT analysis of the Iowa consonant matrices (/p, b, m, f, v, ʃ, dʒ, s, z, t, d, n, g, k/) are presented in Fig. 12. The open bars show IT scores for the vision only condition, the bars with diagonal lines show the scores for the CA processor plus vision condition, and the solid bars show the scores for the IP processor plus vision condition. Note that the viseme and place features are transmitted equally well for all three conditions. The high score for place in the vision only condition is indicative of the high redundancy between assignments for the place and viseme features. That is, a front (bilabial and labiodental) place of articulation usually can be distinguished from other places of articulation through speechreading alone [Owens and Blazek, 1985], and this ability is reflected in the choices for the viseme groupings. Thus, if subjects can distinguish the groups /p, b, m, f, v/, /ʃ, dʒ/, and /d, s, z, t, d, n, g, k/ through speechreading, then the scores for both viseme and place will be high.

Other features that exhibit some redundancy with the viseme groupings are duration and frication. The relatively high scores for these features with vision alone reflect this overlap. On the other hand, the scores for voicing, nasality, and envelope are all low for the vision only condition. These features are invisible on the lips and have little or no redundancy with the viseme groupings.

The scores for both processor plus vision conditions demonstrate increases over the scores for the vision only condition. Especially large increases are found for the features of voicing, duration, and envelope. In addition, the scores for overall information transfer are higher for the processor plus vision conditions.

Comparison of the scores obtained with the two processors indicates superiority of the IP processor for all features except place and viseme, where the scores are about the same. Scores for the IP processor are much higher for the features of voicing, nasality, and envelope. Inasmuch as these are the very features that have little or no redundancy with the viseme groupings, the IP processor might be expected to provide a highly effective supplement to speechreading.

The general finding of superior performance with the IP processor also is evident in the results from IT analysis of the RTI consonant matrices (/ð, s, z, t, d, n, k, l/). Results for the vision only and processor plus vision conditions are presented in Fig. 13, and results for the processor only conditions are presented in Fig. 14. In Fig. 13 the open, diagonally lined and solid bars again show IT scores for vision only, CA processor plus vision and IP processor plus vision, respectively. In Fig. 14 the stippled bars show IT scores for the CA processor only, and the vertically lined bars show the scores for the IP processor only.

For the conditions with a visual component (Fig. 13), high scores again are obtained for the viseme feature. Because the consonants in the RTI test all have a nonlabial place of articulation, however, high scores for the viseme feature merely show that the groups /s, z, t, d, n/, /k, l/, and /ð/ can be distinguished. /ð/ and /l/ usually are visible through tongue protrusion and tongue flap, respectively, even though they have nonlabial places of articulation. Perception of these cues for /ð/ and /l/ can produce relatively high viseme scores for the consonants in the RTI test.

Another effect of the choice of consonants for the RTI test is to hold place of articulation essentially constant. All consonants except /k/ have a mid place of articulation [Singh and Black, 1966]. Thus, the only distinction that has to be made to produce high place scores is the one between /k/ (back place of articulation) and the remaining consonants. The low place score for the vision only condition in Fig. 13 reflects the fact that the place and viseme features are not redundant for the particular consonants of the RTI test. The scores for all other features (voicing, nasality, duration, frication, and envelope) are generally consistent with the scores for the vision only condition of the Iowa test.

Comparison of results across conditions again shows increases over the vision only scores when either processor is used with speechreading. The largest increases are found for the features of voicing, nasality, place, and envelope. The increases for voicing and envelope are quite similar to those found for the Iowa test. The increases for nasality and place, however, are not seen (place) or not as large (nasality) in the Iowa results. The difference in the increases for place can be attributed to the particular choice of consonants in the RTI test, as outlined above. The difference in

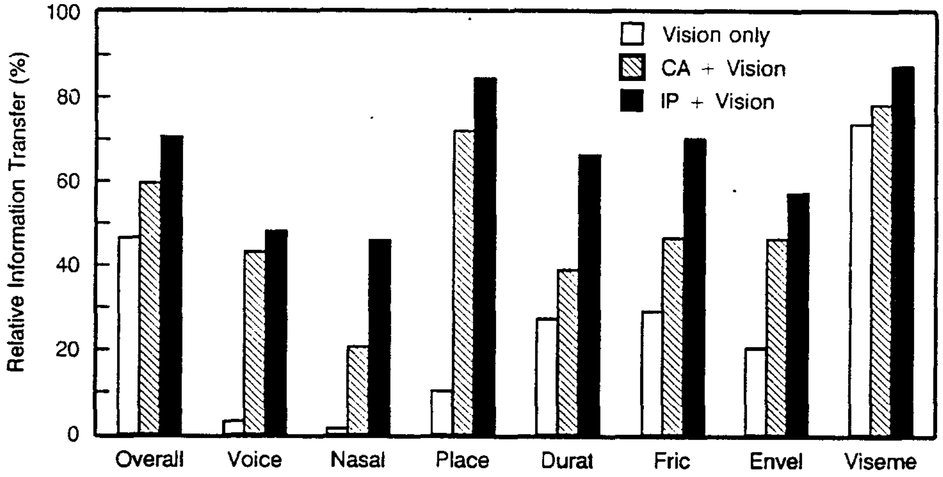


Fig. 13. Relative information transfer of speech features for the vision-only and vision-plus-processor conditions of the RTI consonant test.

the increases for nasality is one of degree in that increases are found for both tests, but the relative increase for the CA processor plus vision over vision only is not as large for the Iowa test compared with the increase for the RTI test. This difference between tests again might be a consequence of the different choices of consonants: the only nasal in the RTI test is /n/, while the Iowa test contains /n/ and /m/. The single distinction of /n/ from the remaining consonants may allow higher scores for the nasal feature in the RTI test.

As with the Iowa test, large increases are found in feature transmission scores when the IP processor is used instead of the CA processor for the vision plus processor conditions. The IP processor produces at least some increase in the score for every studied feature, and substantial increases are demonstrated for the features of nasality, duration, and frication. The same pattern of increases is evident in the scores for the Iowa test; however, the relative increases for the voicing and envelope features are greater with the Iowa test, while the relative increases for the duration and frication features are greater with the RTI test. These differences probably can be attributed to the differences in the consonant sets and to test variability. In all, the patterns of results from the Iowa and RTI tests are remarkably consistent. Both patterns demonstrate substantial gains over vision alone when either processor is used in conjunction with speechreading, and both patterns show superiority of the IP processor. In addition, the particular differences in feature scores found between conditions for one of the tests usually are found for the other test as well.

The results from the RTI test for the processor only conditions (Fig. 14) mirror those reviewed above for the processor plus vision conditions (Fig. 13). Specifically, the IP processor again produces an increase in the score for every studied feature, and substantial increases are found for the features of nasality, duration and frication. Moreover, for all features the

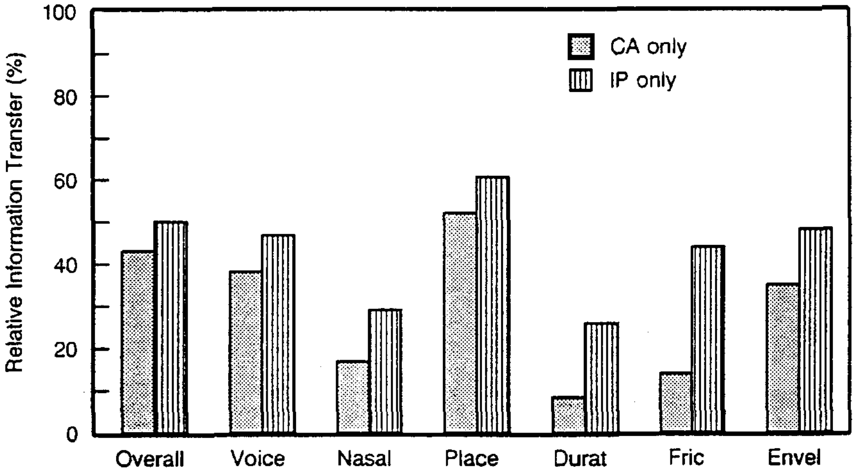


Fig. 14. Relative information transfer of speech features for the processor-only conditions of the RTI consonant test.

ratios of the scores for the CA processor plus vision and IP processor plus vision conditions (Fig. 13) closely approximate the ratios for the CA processor only and IP processor only conditions (Fig. 14). These findings suggest that the IP processor provides additional cues which are utilized by the subjects in both the hearing only and hearing plus vision conditions.

In contrast to the results from the Iowa and RTI consonant tests, the IT scores from the RTI vowel test (/i, I, ɔ, o, u/) indicate superiority of the CA processor. These scores for the vowel test are presented in Fig. 15, where the coding of the bars for the various conditions is identical to the coding used in Figs. 12-14. Comparison of the IT scores between processors shows that the CA processor produces higher or equivalent scores for every feature. For the processor plus vision conditions higher scores are obtained for overall transmission, F1, and duration; and for the processor only conditions higher scores are obtained for these features and F2. In the cases where equivalent scores are found (F2 and viseme features, processor plus vision conditions), ceiling effects may have masked true differences between the processors. A more difficult test (with, for example, more vowels and less redundancy between assignments for the F2 and viseme features) would provide a more sensitive detector of any difference between processors. In any event, the present results show that the CA processor is superior at least for the transmission of F1 and duration information.

The most general observations from the IT data reviewed above are that (a) the IP processor produces higher or essentially equivalent scores for every studied feature of the phonemes in the Iowa and RTI consonant tests and (b) the opposite is found for every studied feature of the phonemes in the RTI vowel test. To evaluate the significance of these observations, a blocked ANOVA was conducted for each test with the features as blocks and the processor plus vision and processor only conditions as the second factor. Because the objective was to compare the contributions of the two processors,

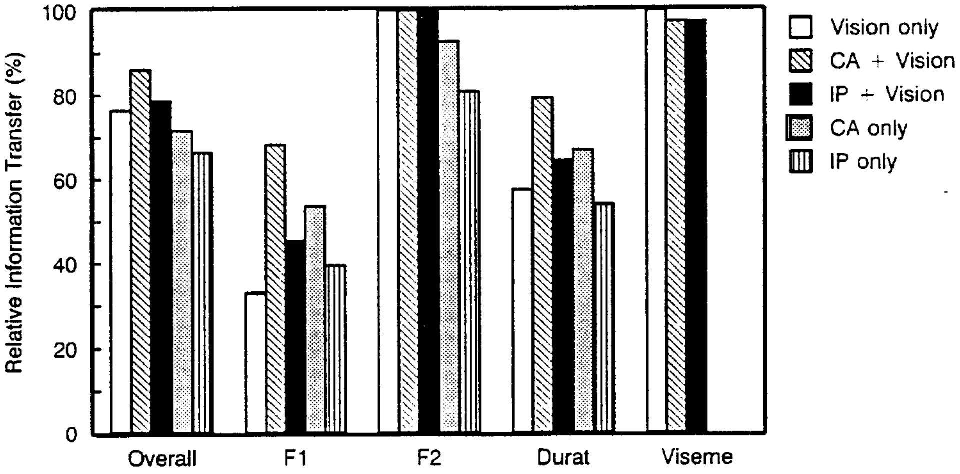


Fig. 15. Relative information transfer of speech features for the RTI vowel test.

the viseme feature was not included in the analysis. The ANOVA results supported the general observations for each of the three tests. In particular, significant effects of conditions were found for all three tests (for the Iowa consonant test $F(1,6) = 18.19$ and $p < 0.01$; for the RTI consonant test $F(2,18) = 19.92$ and $p < 0.001$; and for the RTI vowel test $F(3,9) = 18.33$ and $p < 0.001$). The significant effect for the Iowa test demonstrated superior performance of the IP processor ($p < 0.01$). Post hoc comparison of the means for the four conditions of the RTI consonant test (using the Tukey procedure) also demonstrated superiority of the IP processor (for the processor plus vision conditions $p < 0.01$, and for the processor only conditions $p < 0.05$). Finally, post hoc comparison of the means for the RTI vowel test demonstrated superiority of the CA processor (for both sets of processor conditions, with and without vision, $p < 0.05$).

Representations of Speech Features with CA and IP Processors

The differences in results for the consonant and vowel tests are consistent with differences in the ways in which each processor presents speech stimuli. The CA processor simultaneously presents continuous analog waveforms to all of the stimulation channels in a multielectrode array, while the IP processor presents nonsimultaneous pulses to the same channels. In addition, the type of IP processor used in the present study provides explicit coding of fundamental frequency (F0) and voiced/unvoiced intervals. F0 is coded by initiating stimulation cycles in synchrony with the detected F0 during voiced intervals, and unvoiced intervals are signaled by initiating stimulation cycles at randomly varied times or at rates above the normal range of F0.

A possible advantage of the CA processor is in the preservation of details in the stimulus waveforms; indeed, the results from the vowel test are consistent with the observation that at least some patients can discriminate frequency changes through the range of F1 with CA stimuli. This ability would explain the high IT scores obtained with the CA processor for the F1 feature. Because waveform details are discarded in the IP processor, one might also expect lower scores for F1 with that processor. Differences in scores for transmission of F1 information thus may be attributed to the presence (CA processor) or absence (IP processor) of details in analog waveforms.

Improved perception of F1 also could explain the high IT scores found for the duration feature with the CA processor. The roles of the F1 and duration features are highly correlated for the vowels in the RTI test. This leads to assignment redundancies, that could produce high IT scores for duration even if duration *per se* were not well transmitted. The present results do not allow discrimination among three possibilities: (a) excellent transmission of F1 information only, (b) excellent transmission of duration information only, or (c) excellent transmission of both. Results from previous studies, on the perception of F1 information with CA waveforms [Eddington, 1983; Hochmair-Desoyer and Burian, 1985; White, 1983], support possibilities (a) and (c).

The chief advantage of the IP processor may be the reduction of channel interactions through the use of nonsimultaneous stimuli. This reduction may increase the salience of channel-related cues and therefore improve the representation of frequency components across the speech spectrum. An increased salience of channel cues is likely to become progressively more important as the number of stimulation channels is increased. Thus, this advantage of the IP processor may be more evident in subjects with many channels (e.g., 6 - 10) than in subjects with relatively few channels (e.g., 2 - 4). Indeed, results from previous studies conducted by our group, using a subject fitted with a percutaneous cable, demonstrated large increases in both consonant and vowel identifications when the number of stimulation channels in an IP processor was increased from 4 to 6 [section II; Wilson et al., 1988a].

For the subjects of the present study, with only 2 - 4 channels of stimulation, the representation of frequency components with channel cues is necessarily coarse. However, even a coarse representation may provide a significant advantage for the perception of overall spectral shape. This advantage might be reflected in the improved IT scores for many of the consonant features, especially nasality and voicing. The nasals are characterized by an intense formant at approximately 250 Hz and a relatively weak set of higher formants above 800 Hz. The spectral region between the first and second formants is almost devoid of energy. Recognition of this spectral shape could allow identification of the nasals as a class in the consonant identification tests.

A superior representation of spectral shape also would provide greater access to voicing information. An important cue to voicing is the ratio of energy at low frequencies (e.g., below 1500 Hz) to energy at high frequencies (e.g., above 1500 Hz). A high ratio typically is found for voiced intervals of speech, and a low ratio typically is found for unvoiced intervals. Faithful transmission of the relative energies in these bands would help to produce correct voiced/unvoiced decisions.

The voiced/voiceless distinction might also be aided by the explicit coding of voiced and unvoiced intervals used in the IP processor. Indeed, percepts produced during unvoiced intervals -- with stimulus pulses presented at randomly varied times or at rates above the normal range of F0 -- should be easily discriminated from the percepts produced during voiced intervals -- with stimulus pulses presented at the F0 rate (see, e.g., Dobie and Dillier, 1985; Pfingst, 1985; Shannon, 1983). Such coding of voiced and unvoiced intervals complements the representation of spectral balance between low and high bands. This multiplicity of cues to voicing could increase the subjects' access to voicing information. Improved transmission of voicing information would be expected to produce increases in the IT scores for voicing, envelope, and frication.

A final difference that may favor the IP processor involves the stimuli used for the high frequency (basal) channels. The presence of speech energy at high frequencies is signaled in the CA processor by the delivery of high-frequency analog stimuli to these channels. In contrast, low-frequency pulse trains are used in the IP processor for all channels, avoiding a potential problem with adaptation. Results from both psychophysical [Shannon, 1983] and single-unit [Javel, in press; Moxon, 1971; Parkins, 1986; van den Honert and Stypulkowski, 1987] studies have demonstrated strong adaptation to stimuli with frequencies much above 300 Hz. Thus, the use of such stimuli in the CA processor could produce substantial temporal distortions in the representation of sustained high-frequency components in speech, a problem that may be avoided through the use of an IP processor. A faithful representation of sustained high-frequency components would allow discrimination of the long-duration consonants /ʃ, s, z/ from the remaining consonants in the Iowa and RTI tests. This discrimination would improve the IT score for the duration feature. The IT scores for the envelope and frication features may also be enhanced by a representation that maintains perception of sustained high-frequency sounds.

To summarize the observations on the consonant and vowel results, we note that (a) waveform cues may be used to perceive the F1 feature of vowels with the CA processor and (b) reduced channel interactions and explicit coding of voice/unvoiced information may produce increased scores for the voicing, nasality, frication and envelope features of consonants with the IP processor. In addition, the use of low-frequency pulse trains for the stimulation of basal channels in the IP processor may reduce or eliminate possible deleterious effects of adaptation. This last advantage of IP processors would be expected to produce increases in the transmission of duration, frication, and envelope information for consonants.

investigation of the dependence of pitch on the intensity, frequency and site of intracochlear stimulation. We also plan to investigate in greater detail the large differences in the patterns of pitch percepts obtained under conditions of monopolar and radial bipolar stimulation. These latter studies will of course require subjects implanted with devices that allow such manipulations in the electrode coupling configuration.

V. Psychophysical Studies

The psychophysical studies conducted to date have been aimed at (a) obtaining threshold and loudness data required for the specification of speech processor parameters, (b) obtaining indirect indications of the extent and pattern of nerve survival in the implanted ear, and (c) characterizing patterns of pitch and loudness perception with monopolar and radial bipolar configurations of intracochlear electrodes. The studies to obtain threshold and loudness data for specification of processor parameters have been conducted with all patients in our series, and the studies to obtain indirect indications of nerve survival have been limited mainly to patients with percutaneous access to their implanted electrodes. These latter studies have included measures of (a) simultaneous and nonsimultaneous channel interactions; (b) thresholds to low frequency (e.g., 100 Hz) and high frequency (e.g., 1000 Hz) sinusoids; (c) dynamic ranges between threshold and uncomfortable loudness levels for sinusoidal and pulsatile stimuli; and (d) differences in thresholds for monopolar and radial bipolar stimulation with the UCSF electrode array.

The studies to characterize patterns of pitch and loudness perception with different electrode coupling configurations were conducted with one of our percutaneous cable patients (MH). The results demonstrated some striking differences in the patterns of pitch percepts reported by this patient under conditions of monopolar and radial bipolar stimulation with the UCSF electrode array. In particular, the repetition frequency of pulsatile stimulation was a much more salient cue for pitch with the monopolar coupling configuration than with the bipolar configuration, while the site of intracochlear stimulation generally had only minor effects on pitch judgments. The latter cue did affect pitch judgments -- in a complex way -- for selected channels among the bipolar electrodes. Judgments of pitch also were strongly dependent on the *intensity* of stimulation for certain electrode channels and stimulation frequencies for both coupling configurations. In fact, the intensity of stimulation in many cases was a more salient cue for pitch than either the stimulation frequency or the site at which the stimuli were delivered. In all such cases increases in stimulus intensity produced increases in perceived pitch. This "covariance" between intensity (or the percept of loudness) and pitch has been noted by other investigators (e.g., Simmons et al., 1979). It presents an obvious problem for speech processors that attempt to represent the frequency content of an acoustic input by the frequency and/or site of intracochlear stimulation. Clearly, this representation will be distorted for patients who perceive large changes in pitch when the intensity of stimulation is varied.

A complete discussion of the implications of these findings for processor design is presented in the fifth quarterly progress report for this project [Wilson et al., 1986b]. Details of the experimental methods and all results also are presented in that report. Results from preliminary studies with other patients suggest that the apparent dependence of pitch on intensity of stimulation is highly variable across patients (and across electrode channels within patients). Indeed, some patients have moderate *decreases* in pitch with increases in intensity and some patients have stable pitch percepts over the entire dynamic range of intensities. We plan a systematic

VI. Development of a Next-Generation Auditory Prosthesis

We have been collaborating with UCSF in the development of a next-generation auditory prosthesis. In particular, we have played major roles in the specification and design of the speech processor, transcutaneous transmission system (TTS), and clinical fitting system for this new device. The design of these components is based on the results from our studies comparing analog and pulsatile coding strategies in tests with individual implant patients (see section II of this report). A major conclusion from those studies is that substantial gains in speech understanding can be made for implant recipients by (a) selecting the best type of speech processor for each patient and (b) using implanted and external hardware capable of supporting a wide range of different processing strategies.

The new prosthesis, to be produced by MiniMed Technologies, will be capable of supporting both the CA and IP processing strategies and also will have a TTS that provides up to eight channels of current-controlled outputs for intracochlear stimulation. In fact, the TTS will have unprecedented flexibility for such a system. The current drivers are fully isolated and can be individually programmed to provide outputs in 3 percent increments between 2 and 2000 µA. The drivers also can be instructed to provide a zero output or negative outputs (the direction of current flow is controlled by a switch matrix). All eight driver outputs can be simultaneously updated every 77 µs (13 kHz sampling rate) and a smaller number of driver outputs can be updated at higher rates. For example, one channel can be updated every 12.3 µs (81 kHz sampling rate). This feature permits a mixture of sampling rates for the various stimulation channels which might be most useful for the generation of stimuli for CA processors. In particular, high frequency analog stimuli can be "reconstructed" at high sampling rates for basal electrode channels while low frequency stimuli can be reconstructed at low sampling rates for apical channels. Finally, the outputs of the current drivers can be connected (under external control) to form different electrode coupling configurations. Both the monopolar and radial bipolar configurations will be supported, as will "hybrid" mixtures of the two. In addition, a longitudinal bipolar configuration can be implemented if nonsimultaneous stimuli are used. All outputs from the current drivers are capacitively coupled to the selected intracochlear electrodes to ensure charge balancing of the stimuli.

The significance of the new UCSF device for further development of cochlear prostheses is that investigators will have access to patients with a highly-transparent TTS. This will allow evaluation of an extremely wide range of processing strategies and electrode coupling configurations in tests with individual patients without having to use a percutaneous cable. Thus, we will be able to retain most of the flexibility provided by the cable while eliminating the potential complications of cable use.

VII. Concluding Remarks

We have developed and evaluated two new classes of speech processors, one of which appears to be highly useful for patients with psychophysical signs of poor nerve survival (the IP processor) and the other of which appears to be highly appropriate for use in cochlear prostheses or implant patients with only two or three functional channels of stimulation (the B/P processor).

Findings from the evaluation studies have led to the development of a portable, real-time processor and to the development of a next-generation cochlear prosthesis.

In future work we plan to continue evaluation of the IP, B/P and other processing strategies in tests with much larger numbers of patients, implanted with a wide variety of intracochlear electrodes (e.g., MiniMed, Nucleus, and Symbion). The additional processing strategies will include the hybrid CA/IP strategy mentioned in sections II and IV of this report; promising variations of the basic IP processor; and variations of the CA processor that use cross coupling of outputs to control compression in each stimulation channel [White, 1986]. Finally, we will continue work to (a) determine how the electrode coupling configuration affects the perception of loudness and pitch with cochlear implants; (b) compare indirect psychophysical measures of nerve survival in the implanted cochlea with measures of speech perception; (c) develop and apply more realistic models of neural responsiveness to intracochlear electrical stimulation; and (d) design and fabricate further generations of portable, real-time speech processors to implement the best strategies identified in the evaluation studies.

VIII. Acknowledgments

We thank the patients who participated in the evaluation studies for their dedicated effort and pioneering spirit. We are pleased to acknowledge the important scientific contributions of L.J. Dent, D.K. Eddington, J.C. Farmer, Jr., F.T. Hambrecht, W.J. Heetderks, P.D. Kenan, D.K. Kessler, G.E. Loeb, M.M. Merzenich, J.F. Patrick, W.M. Rabinowitz, S.J. Rebscher, R.A. Schindler, R.V. Shannon, M.W. Skinner, S.D. Soli, L. Vurek, B.A. Weber, M.W. White, D. Wilkinson, and R.D. Wolford.

IX. References

Breeuwer M, Plomp R (1984) Speechreading supplemented with frequency-selective sound-pressure information. J. Acoust. Soc. Am. 76: 686-691.

Brimacombe JA, Webb RL, Dowell RC, Mecklenburg DJ, Beiter AL, Barker MJ, Clark GM (1988) Speech recognition abilities in profoundly deafened adults using the Nucleus 22 channel cochlear implant system. In: Banfai P (ed) Cochlear implant: Current situation. Erkelenz, West Germany: Rudolf Bermann GmbH, pp. 487-490.

Clark GM, Blamey PJ, Brown AM, Gusby PA, Dowell RC, Franz BK-H, Pyman BC, Shepherd RK, Tong YC, Webb RL, Hirshorn MS, Kuzma J, Mecklenburg DJ, Money DK, Patrick JF, Seligman PM (1987) The University of Melbourne-Nucleus multi-electrode cochlear implant. Basel: Karger.

Danley MJ, Fretz RJ (1982) Design and functioning of the single-electrode cochlear implant. Ann. Otol. Rhinol. Laryngol. 91, Suppl. 91: 21-26.

De Filippo CL, Scott BL (1978) A method for training and evaluating the reception of ongoing speech. J. Acoust. Soc. Am. 63: 1186-1192.

Dobie RA, Dillier N (1985) Some aspects of temporal coding for single- channel electrical stimulation of the cochlea. Hearing Res. 18: 41-55.

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.

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

Eisenberg LS, Maltan AA, Portillo F, Mobley JP, House WF (1987) Electrical stimulation of the auditory brain stem structure in deafened adults. J. Rehab. Res. & Devel. 24: 9-22.

Finley CC, Wilson BS, Lawson DT (1987) Speech processors for auditory prostheses. Seventh Quarterly Progress Report, NIH project N01-NS-5-2396. Bethesda, MD: Neural Prosthesis Program, National Institutes of Health.

Gantz BJ (1987) Cochlear implants: An overview. Adv. Otolaryngol. Head & Neck Surg. 1: 171-200.

Gantz BJ, Tyler RS, Knutson JF, Woodworth G, Abbas P, McCabe BF, Hinrichs J, Tye-Murray N, Lansing C, Kuk F, Brown C (1988) Evaluation of five different cochlear implant designs: Audiologic assessment and predictors of performance. Laryngoscope 98: 1100-1106.

Gardi JN (1985) Human brain stem and middle latency responses to electrical stimulation: Preliminary observations. In: Schindler RA, Merzenich MM (eds) Cochlear implants. New York: Raven Press, pp. 351-363.

Grether CB (1970) Psychoacoustic assessment of speech communication systems: The diagnostic discrimination test. Project Themis. Air Force Office of Scientific Research, contract F44620-69-C-0033.

Hochmair ES, Hochmair-Desoyer IJ (1985) Aspects of sound signal processing using the Vienna intra- and extracochlear implants. In: Schindler RA, Merzenich MM (eds) Cochlear implants. New York: Raven Press, pp. 101-110.

Hochmair-Desoyer IJ, Burian K (1985) Reimplantation of a molded scala tympani electrode: Impact on psychophysical and speech discrimination abilities. Ann. Otol. Rhinol. Laryngol. 94: 65-70.

Hochmair-Desoyer IJ, Hochmair ES, Stiglbrunner HK (1985) Psychoacoustic temporal processing and speech understanding in cochlear implant patients. In: Schindler RA, Merzenich MM (eds) Cochlear implants. New York: Raven Press, pp. 291-304.

Hollander M, Wolfe DA (1973) Nonparametric statistical methods. New York: John Wiley & Sons, pp. 26-33.

Javel E (in press) Acoustic and electrical encoding of temporal information. To appear in Miller JM, Spelman FA (eds), Models of the electrically stimulated cochlea, Springer-Verlag.

Lawson DT, Wilson BS, Finley CC (1989) Speech processors for auditory prostheses. Fourteenth Quarterly Progress Report, NIH project N01-NS-5-2396. Bethesda, MD: Neural Prosthesis Program, National Institutes of Health.

Loeb GE, Byers CL, Rebscher SJ, Casey DE, Fong MM, Schindler RA, Gray RF, Merzenich MM (1983) Design and fabrication of an experimental cochlear prosthesis. Med. & Biol. Eng. & Comput. 21: 241-254.

Merzenich MM (1985) UCSF cochlear implant device. In: Schindler RA, Merzenich MM (eds) Cochlear implants. New York: Raven Press, pp. 121-129.

Merzenich MM, Leake-Jones P, Vivion M, White M, Silverman M (1978) Development of multichannel electrodes for an auditory prosthesis. Fourth Quarterly Progress Report, NIH project N01-NS-7-2367. Bethesda, MD: Neural Prosthesis Program, National Institutes of Health.

Millar JB, Tong YC, Clark GM (1984) Speech processing for cochlear implant prostheses. J. Speech & Hear. Res. 27: 280-296.

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

Moore BCJ (1985) Speech coding for cochlear implants. In: Grey RF (ed) Cochlear implants. San Diego, CA: College-Hill Press, pp. 163-179.

Moxon EC (1971) Neural and mechanical responses to electric stimulation of the cat's inner ear. Doctoral dissertation, MIT, Cambridge, MA.

Ochs MT, White MW, Merzenich MM, Schubert ED (1985) Speech recognition in single- and multichannel cochlear implants. J. Acoust. Soc. Am. 77, Suppl. 1: S81.

Owens E, Blazek B (1985) Visemes observed by hearing-impaired and normal-hearing adult viewers. J. Speech & Hear. Res. 28: 381-393.

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

Owens E, Raggio M (1987) The UCSF tracking procedure for evaluation and training of speech reception by hearing-impaired adults. J. Speech & Hear. Disorders 52: 120-128.

Parkins CW (1986) Cochlear prostheses. In: Altschuler RA, Hoffman DW, Bobbin RP (eds) Neurobiology of hearing: The cochlea. New York: Raven Press, pp. 455-473.

Pfingst BE (1985) Psychophysical data from cochlear implants: Relevance to strategies for rehabilitation. Seminars in Hearing 6: 7-21.

Pfingst BE (1986) Stimulation and encoding strategies for cochlear prostheses. Otolaryngol. Clin. N. Amer. 19: 219-235.

Pfingst BE, Glass I, Spelman FA, Sutton D (1985) Psychophysical studies of cochlear implants in monkeys: Clinical implications. In: Schindler RA, Merzenich MM (eds) Cochlear implants. New York: Raven Press, pp. 305-321.

Pfingst BE, Sutton D (1983) Relation of cochlear implant function to histopathology in monkeys. Ann. N.Y. Acad. Sci. 405: 224-239.

Schindler RA, Kessler DK (1987) The UCSF/Storz cochlear implant: Patient performance. Am. J. Otol. 8: 247-255.

Schindler RA, Kessler DK, Rebscher SJ, Jackler RK, Merzenich MM (1987) Surgical considerations and hearing results with the UCSF/Storz cochlear implant. Laryngoscope 97: 50-56.

Schindler RA, Kessler DK, Rebscher SJ, Yanda JL, Jackler RK (1986) The UCSF/Storz multichannel cochlear implant: Patient results. Laryngoscope 96: 597-603.

Schubert ED (1985) Some limitations on speech coding for implants. In: Schindler RA, Merzenich MM (eds) Cochlear implants. New York: Raven Press, pp. 269-276.

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

Simmons FB, Walker MG, Mathews RG, White RL (1979) Percepts and discrimination by auditory nerve stimulation: A summary of results and some proposals for nerve viability evaluation. In: McPherson DL, Davis MS (eds) Advances in prosthetic devices for the deaf: A technical workshop. Rochester, New York: National Technical Institute for the Deaf, Rochester Institute of Technology, pp. 271-274.

Soli SD (in press) Perceptual evaluation of a neurally-based encoding strategy for cochlear implants. To appear in Miller JM, Spelman FA (eds), Models of the electrically stimulated cochlea, Springer-Verlag.

Soli SD, Arabie P, Carroll JD (1986) Discrete representation of perceptual structure underlying consonant confusions. J. Acoust. Soc. Am. 79: 826-837.

Soli SD, Wilson BS (1988) Within-subject comparisons of analog and pulsatile speech processors for cochlear implants. J. Acoust. Soc. Am. 84, Suppl. 1: S41.

Singh S, Black JW (1966) Study of twenty-six intervocalic consonants as spoken by four language groups. J. Acoust. Soc. Am. 39: 372-387.

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 (1983) The Iowa cochlear-implant test battery. Laboratory Report, University of Iowa at Iowa City, Department of Otolaryngology - Head and Neck Surgery.

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

Van Tasell DJ, Soli SD, Kirby VM, Widin GP (1987) Speech waveform envelope cues for consonant recognition. J. Acoust. Soc. Am. 82: 1152-1161.

Wang MD, Bilger RC (1973) Consonant confusions in noise: A study of perceptual features. J. Acoust. Soc. Am. 54: 1248-1266.

White M, Ochs M, Raggio M, Morledge D, Merzenich M (1985) Formant discrimination in a multichannel cochlear prosthesis. J. Acoust. Soc. Am. 77, Suppl. 1: S81.

White MW (1983) Formant frequency discrimination and recognition in subjects implanted with intracochlear stimulating electrodes. Ann. N.Y. Acad. Sci. 405: 348-359.

White MW (1986) Compression systems for hearing aids and cochlear prostheses. J. Rehab. Res. & Devel. 23: 25-39.

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

Wilson BS, Finley CC (1985) A computer-based simulator of speech processors for auditory prostheses. ARO Abstracts, 8th Midwinter Res. Conf., St. Petersburg, FL, p. 209.

Wilson BS, Finley CC, Farmer JC Jr, Lawson DT, Weber BA, Wolford RD, Kenan PD, White MW, Merzenich MM, Schindler RA (1988a) Comparative studies of speech processing strategies for cochlear implants. Laryngoscope 98: 1069-1077.

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

Wilson BS, Finley CC, Lawson DT (1986a) Speech processors for auditory prostheses. Second Quarterly Progress Report, NIH project N01-NS-5-2396. Bethesda, MD: Neural Prosthesis Program, National Institutes of Health.

Wilson BS, Finley CC, Lawson DT (1986b) Speech processors for auditory prostheses. Fifth Quarterly Progress Report, NIH project N01-NS-5-2396. Bethesda, MD: Neural Prosthesis Program, National Institutes of Health.

Wilson BS, Finley CC, Lawson DT (1987) Speech processors for auditory prostheses. Eighth Quarterly Progress Report, NIH project N01-NS-5-2396. Bethesda, MD: Neural Prosthesis Program, National Institutes of Health.

Wilson BS, Finley CC, Lawson DT (in press) Representations of speech features with cochlear implants. To appear in Miller JM, Spelman FA (eds), Models of the electrically stimulated cochlea, Springer-Verlag.

Wilson BS, Finley CC, Lawson DT, Wolford RD (1988b) Speech processors for cochlear prostheses. Proc. IEEE 76: 1143-1154.

Wilson BS, Finley CC, Lawson DT, Wolford RD (1988c) Direct comparisons of analog and pulsatile coding strategies with six cochlear implant patients. In prep.

Wilson BS, Schindler RA, Finley CC, Kessler DK, Lawson DT, Wolford RD (1988d) Present status and future enhancements of the UCSF cochlear prosthesis. In: Banfai P (ed) Cochlear implant: Current situation. Erkelenz, West Germany: Rudolf Bermann GmbH, pp. 395-427.

Appendix 1: Key Contents of Quarterly Progress Reports

Report	Key contents
1	Review of psychophysical and speech perception tests with subject SG
2	Evaluation of 13 major classes of speech processing strategies in tests with subject MH; further development of an interleaved pulses (IP) processor
3	Report on the development of real-time, portable speech processors; presentation of preliminary results from measurements of intracochlear electric field patterns using a percutaneous cable
4	Evaluation of idealized implementations of the processing strategy used in the Nucleus auditory prosthesis
5	Report on studies of loudness and pitch perception under conditions of stimulation with monopolar and radial bipolar electrodes
6	Direct comparisons of analog and pulsatile coding strategies in studies with six patients implanted with the UCSF/Storz electrode array; further development and application of a portable speech processor
7	Detailed description of a portable speech processor for multichannel cochlear prostheses
8	Report on the design and-evaluation of a two channel, "Breeuwer/Plomp" processor
9	Detailed report of comparisons of analog and pulsatile coding strategies for multichannel cochlear prostheses
10	Review of recent results from the clinical trials of the UCSF/Storz cochlear prosthesis
11	Report on work to upgrade the capabilities of the Cochlear Implant Laboratory at Duke Medical Center; report on the collaborative development of a next-generation auditory prosthesis
12	Detailed report on representations of speech features with cochlear implants
13	Detailed report on models of neural responsiveness to intracochlear electrical stimulation
14	Report on binary comparisons (e.g., processor A *versus* processor B, male *versus* female, subject A *versus* subject B) of processor performance

Appendix 2: Publications and Presentations for NIH project N01-NS-5-2396

Publications

Wilson, B.S., C.C. Finley and D.T. Lawson: Representations of speech features with cochlear implants. In J.M. Miller and F.A. Spelman (Eds.), Models of the Electrically Stimulated Cochlea, Springer-Verlag, in press.

Finley, C.C., B.S. Wilson and M.W. White: Models of neural responsiveness to electrical stimulation. In J.M. Miller and F.A. Spelman (Eds.), Models of the Electrically Stimulated Cochlea, Springer-Verlag, in press.

Finley, C.C.: A finite-element model of radial 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, in press.

Wilson, B.S., C.C. Finley, D.T. Lawson and R.D. Wolford: Speech processors for cochlear prostheses. In the special issue of Proc. IEEE on "Emerging Electromedical Systems," vol. 76: 1143-1154, 1988.

Wilson, B.S., R.A. Schindler, C.C. Finley, D.K. Kessler, D.T. Lawson and R.D. Wolford: Present status and future enhancements of the UCSF cochlear prosthesis. In P. Banfai (Ed.), Cochlear Implants: Current Situation, Rudolf Bermann GmbH, Erkelenz, West Germany, 1988, pp. 395-427.

Wilson, B.S. (moderator), L.J. Dent, N. Dillier, D.K. Eddington, I.J. Hochmair-Desoyer, B.E. Pfingst, J. Patrick, W. Sürth and J. Walliker (panelists): Round table discussion on speech coding. In P. Banfai (Ed.), Cochlear Implants: Current Situation, Rudolf Bermann GmbH, Erkelenz, West Germany, 1988, pp. 693-704.

Wilson, B.S., C.C. Finley, J.C. Farmer, Jr., D.T. Lawson, B.A. Weber, R.D. Wolford, P.D. Kenan, M.W. White, M.M. Merzenich and R.A. Schindler: Comparative studies of speech processing strategies for cochlear implants. Laryngoscope, 98: 1069-1077, 1988.

Wilson, B.S., C.C. Finley, M.W. White and D.T. Lawson: Comparisons of processing strategies for multichannel auditory prostheses. In Proc. Ninth Ann. Conf. Engineering in Medicine and Biology Soc., IEEE Press, 1987, pp. 1908-1910.

White, M.W., C.C. Finley and B.S. Wilson: Electrical stimulation model of the auditory nerve: Stochastic response characteristics. In Proc. Ninth Ann. Conf. Engineering in Medicine and Biology Soc., IEEE Press, 1987, pp. 1906-1907.

Finley, C.C., B.S. Wilson and M.W. White: A finite-element model of bipolar field patterns in the electrically stimulated cochlea -- A two dimensional approximation. In Proc. Ninth Ann. Conf. Engineering in Medicine and Biology Soc., IEEE Press, 1987, pp. 1901-1903.

Abstracts and Presentations

Wilson, B.S.: Speech processors for auditory prostheses. Presented at the Neural Prosthesis Workshop, National Institutes of Health, Bethesda, MD, 1985, 1986, 1987 and 1988.

Wilson, B.S.: Speech processing strategies for cochlear implants. Invited speaker presentation to be given at the University of Iowa, Iowa City, IA, August 28, 1989.

Wilson, B.S.: Within patient evaluation of speech processors. Invited paper to be presented at the Engineering Foundation Conference on Implantable Auditory Prostheses, Potosi, MO, July 30 to August 4, 1989.

Wilson, B.S.: Comparison of encoding schemes. Invited paper presented at the 25th Anniversary Symposium of the Kresge Hearing Research Institute, Ann Arbor, MI, Oct. 3-5, 1988.

Finley, C.C.: 3D finite element analysis. Invited paper presented at the 25th Anniversary Symposium of the Kresge Hearing Research Institute, Ann Arbor, MI, Oct. 3-5, 1988.

Finley, C.C.: Finite element modeling of intracochlear electrodes. Presented at the Neural Prosthesis Workshop, National Institutes of Health, Bethesda, MD, 1988.

Soli, S.D. and B.S. Wilson: Within-subject comparisons of analog and pulsatile speech processors for cochlear implants. J. Acoust. Soc. Am., 84 (Suppl. 1): S41, 1988.

Finley, C.C.: Design considerations for auditory prostheses. Invited paper presented at the World Cong. on Med. Phys. and Biomed. Eng., San Antonio, TX, Aug. 6-12, 1988.

Finley, C.C.: Co-Chairman, Session on Auditory System Research. World Cong. on Med. Phys. and Biomed. Eng., San Antonio, TX, Aug. 6-12, 1988.

Wilson, B.S.: Various coding schemes used. Invited paper presented at the Cochlear Implant Consensus Development Conf., National Institutes of Health, Bethesda, MD, May 2-4, 1988.

Lawson, D.T.: Processing strategies for cochlear implants. Invited faculty lecture, Mayo Symp. on Continuing Education in Audiology, Jacksonville, FL, Feb. 19-20, 1988.

Wilson, B.S.: Review of RTI research on coding strategies for cochlear prostheses. Invited lecture presented at the 3M Company in St. Paul, MN, Nov. 12, 1987.

Wilson, B.S.: Moderator, Speech Coding Panel, International Cochlear Implant Symposium 1987, Düren, West Germany, Sept. 7-11, 1987.

Schindler, R.A. and B.S. Wilson: Present status and future enhancements of the UCSF/RTI/Duke cochlear implant, Abstr. International Cochlear Implant Symposium 1987, Düren, West Germany, Sept. 7-11, 1987, p. 57.

Finley, C.C.: Status of current spread in auditory prostheses. Invited paper presented at the Gordon Research Conference on Implantable Auditory Prostheses, New London, NH, June 29 - July 3, 1987.

Wilson, B.S.: Factors in coding speech for auditory prostheses. Invited paper presented at the Gordon Research Conference on Implantable Auditory Prostheses, New London, NH, June 29 - July 3, 1987.

Finley, C.C.: Electrode design and stimulus shaping. Invited paper presented at the Gordon Research Conference on Implantable Auditory Prostheses, New London, NH, June 29 - July 3, 1987.

Wilson, B.S., C.C. Finley, M.W. White and D.T. Lawson: Comparisons of processing strategies for multichannel auditory prostheses. Invited paper presented in the special session on cochlear implants, 1987 IEEE/EMBS Meeting, Boston, MA, Nov. 13-16, 1987.

White, M.W., C.C. Finley and B.S. Wilson: Electrical stimulation model of the auditory nerve: Stochastic response characteristics. Invited paper presented in the special session on cochlear implants, 1987 IEEE/EMBS Meeting, Boston, MA, Nov. 13-16, 1987.

Finley, C.C., B.S. Wilson and M.W. White: A finite-element model of bipolar field patterns in the electrically stimulated cochlea -- A two dimensional approximation. Invited paper presented in the special session on cochlear implants, 1987 IEEE/EMBS Meeting, Boston, MA, Nov. 13-16, 1987.

Finley, C.C., B.S. Wilson and M.W. White: Models of afferent neurons in the electrically stimulated ear. Invited paper presented in the special session on mathematical modeling related to functional electrical stimulation, 1987 IEEE/EMBS Meeting, Boston, MA, Nov. 13- 16, 1987.

Farmer, J.C., Jr. and B.S. Wilson: Cochlear implantation for the profoundly deaf. Invited seminar presentation, Department of Physiology, Duke University Medical Center, June 18, 1987.

Wilson, B.S.: The RTI/Duke cochlear implant program. Presented to the Executive Committee of the Research Triangle Institute (RTI) Board of Governors, June 17, 1987.

Finley, C.C. and R.D. Wolford: Cochlear implants and the Duke Center for the Severely Hearing Impaired. Invited lecture, Departments of Otolaryngology and Audiology, UNC School of Medicine, Chapel Hill, NC, May 8, 1987.

Wilson, B.S., C.C. Finley, J.C. Farmer, Jr., B.A. Weber, D.T. Lawson, R.D. Wolford, P.D. Kenan, M.W. White, M.M. Merzenich and R.A. Schindler: Comparative studies of speech processing strategies for cochlear implants. Abstracts of the 90th Annual Meeting of the Triological Society, Denver, CO, April 28-30, 1987.

Lawson, D.T.: Cochlear implants. Invited presentation in the UNC-Greensboro series of Psychology Colloquia, April 10, 1987.

Wilson, B.S.: Cochlear implants. Invited lecture presented in the session on auditory signal processing, First North Carolina Workshop on Bioelectronics, Quail Roost, NC, Oct. 24-26, 1986.

Lawson, D.T.: Cochlear implants. Invited luncheon address at the Annual Meeting of the NC Regional chapter of the Acoustical Society of America, Tanglewood Park, NC, Oct. 9-10, 1986.

Wilson, B.S.: Processing strategies for cochlear implants. Invited faculty lecture for the special course on cochlear implants at the Annual Meeting of the American College of Otolaryngologists, San Antonio, TX, Sept. 18-19, 1986.

Kenan, P.D., J.C. Farmer, Jr., B.A. Weber and B.S. Wilson: Cochlear implants. Invited lecture presented at the Annual Meeting of the Mecklenberg County Otolaryngology, Head and Neck Surgery Soc., Charlotte, NC, Fall, 1986.

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

Finley, C.C. and B.S. Wilson: Field patterns in the electrically-stimulated human cochlea, IUPS Satellite Symposium on Advances in Auditory Neuroscience, San Francisco, CA, July 8-11, 1986.

Wilson, B.S.: Coding strategies for cochlear implants. Invited speaker presentation at the Kresge Hearing Research Institute, University of Michigan, Ann Arbor, MI, May 22, 1986.

Wilson, B.S.: Comparison of strategies for coding speech with multichannel auditory prostheses. Invited faculty lecture presented at the Conference on Speech Recognition with Cochlear Implants, New York University, April 17-19, 1986.

Finley, C.C. and B.S. Wilson: Sampling of electric fields by myelinated intracochlear neurons. ARO Abstracts, 9th Midwinter Research Conference, p. 170, 1986.

Wilson, B.S. and C.C. Finley: Latency fields in electrically-evoked hearing. ARO Abstracts, 9th Midwinter Research Conference, pp. 170- 171, 1986.

Farmer, J.C., Jr., P.D. Kenan and B.S. Wilson: Cochlear implants. Presented at Surgical Grand Rounds, Duke University Medical Center, November, 1985.

Finley, C.C. and B.S. Wilson: Models of neural stimulation for electrically-evoked hearing. Invited paper presented in the special session on neurostimulation, ACEMB Meeting, Sept. 30 - Oct. 2, 1985.

Wilson, B.S. and C.C. Finley: Speech processors for auditory prostheses. Invited paper presented in the special session on signal processing for the hearing impaired, IEEE Bioengineering Conference, Sept. 27-30, 1985.

Finley, C.C. and B.S. Wilson: A simple finite-difference model of field patterns produced by bipolar electrodes of the UCSF array. IEEE Bioengineering Conference, Sept. 27-30, 1985.