Revised version – B X

Eighth Quarterly Progress Report

June 27 through September 26, 1987

NIH Contract N01-NS-5-2396

Speech Processors for Auditory Prostheses

Prepared by

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

Neuroscience Program Office

Research Triangle Institute

Research Triangle Park, NC 27709

CONTENTS

I. Introduction

The purpose of this project is to design and evaluate speech processors for 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. Work in the present quarter included the following:

1. Further development of the multichannel speech processor for patient MH, as outlined in our last quarterly progress report;
2. Evaluation of alternative hardware choices and designs to increase the flexibility and reduce the power consumption of this portable processor;
3. Continued psychophysical and speech perception studies with patient MH;
4. Preparation of two invited papers, one for a chapter in Cochlear Implants 1987, to be published by Springer-Verlag, and the other for the upcoming special issue of the Proc. IEEE on "Emerging Electromedical Systems;"
5. Presentation of project results in three invited lectures at the Gordon Research Conference on "Implantable Auditory Prostheses" and in one invited lecture at the International Cochlear Implant Symposium 1987, in Düren, West Germany; and
6. Continued collaboration with the UCSF team in the development of the speech processor and transcutaneous transmission system for a next-generation auditory prosthesis.

In this report we will describe one aspect of our studies with patients implanted with the 4-channel UCSF/Storz auditory prosthesis. As outlined in QPR 6, the main purpose of these studies was to compare the performance of each patient's compressed analog (CA) processor with the performance of the interleaved pulses (IP) processors developed in the early stages of this project. A secondary but nevertheless important goal of the studies with these patients was to evaluate a new two-channel processing strategy for cochlear implants. This evaluation was motivated by the need to develop good strategies for 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; 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 or partial insertion of the electrode array.

The major topic of this report is the evaluation of the new two-channel processing strategy. A detailed description of the studies to compare the CA and IP strategies will be presented in our next progress report for this project.

II. Evaluation of Two Channel, "Breeuwer/Plomp" Processors for Cochlear Implants

A. Introduction

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. Breeuwer and Plomp suggest in their paper that the excellent results obtained with such a simple acoustic supplement might be explained by (a) the perceived ratio between high-band and low-band energies providing good coding of the important distinctions between voiced and unvoiced intervals in connected speech, and (b) the perceived overall amplitude of the combined filter outputs providing good coding of the temporal dynamics (prosodic features) of speech. The encouraging results obtained with normal-hearing subjects led Breeuwer and Plomp to suggest that this two channel strategy might be useful in prostheses for the deaf.

In the present report we describe our evaluation of the "Breeuwer/Plomp" strategy in tests with cochlear implant patients. A prime motivation for this study was to determine the efficacy of the Breeuwer/Plomp 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; 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 or partial insertion of the electrode array.

B. Methods

Subjects

Six cochlear implant patients participated in this study as part of an extensive series of tests to compare alternative processing strategies for auditory prostheses (Wilson et al., 1986 and 1987a,b). All six had been implanted with the multichannel electrode array developed at the University of California at San Francisco (UCSF) and subsequently manufactured by Storz Instrument Company of St. Louis. This array has eight pairs of bipolar electrodes, with a 2 mm spacing between pairs (Loeb et al., 1983). The electrode array is inserted into the scala tympani through an opening made at the round window (Schindler et al., 1986 and 1987). The maximum depth of insertion is 25 mm. In ears with good nerve survival such an array would be expected to allow a high degree of spatial selectivity in the excitation of auditory neurons (Merzenich and White, 1977; van den Honert and Stypulkowski, 1987).

The first patient was fitted with a percutaneous cable and the remaining five with the four-channel transcutaneous transmission system (TTS) of the USCF/Storz prosthesis (Schindler _et al_., 1986). 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 pairs of bipolar electrodes 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 frequency bands of the TTS channels. The principal limitations of the TTS for implementing optimized versions of the processors described in this paper are (a) inadequate levels of voltage compliance for stimulation with short-duration pulses and (b) lack of current control in the stimulus waveforms. The implications of these limitations on processor design and performance are discussed in detail elsewhere (Wilson _et al_., 1987b).

Processors

Our application of the processing strategy described by Breeuwer and Plomp consisted of mapping the root-mean-square (RMS) energies of the two octave-band filters onto the dynamic range of electrically evoked hearing for two channels of intracochlear stimulation. The channels chosen for each subject provided distinct "place pitch" percepts. That is, differences in timbre or pitch for loudness-balanced stimuli allowed excellent discrimination of the selected channels for all subjects.

In our implementation of the processor shown in Fig. 1 the two bandpass filters have corner frequencies of 364 and 707 Hz and 2235 and 4470 Hz respectively. Fourth-order Butterworth filters are used (2 poles per skirt). The RMS energy in each band is sensed by a full-wave rectifier and a series low-pass filter at the output of each bandpass filter. A "post

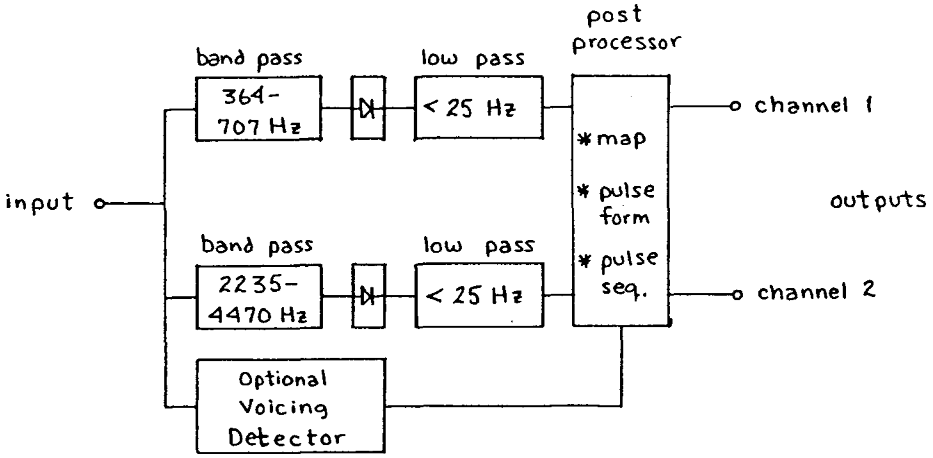


Fig. 1. Block diagram of a two channel, "Breeuwer/Plomp" processor.

processor" monitors these two RMS outputs continually, coding each for stimulation of its assigned electrode pair only if the RMS energy is above a preset "noise threshold." The amplitudes of pulsatile stimuli delivered to the electrodes are calculated from a logarithmic mapping law of the form:

$$\text{pulse amplitude} = A \times \log(\text{RMS level}) + k,$$

where parameters "A" and "k" are determined for each channel on the basis of threshold and most-comfortable loudness level measurements. Finally, the voicing detector senses the fundamental frequency of voiced speech sounds and whether a given speech input is voiced (periodic) or unvoiced (aperiodic). The output of the voicing detector optionally can be used by the post processor to control the timing of update cycles, as described below. Pulsatile stimuli always are presented nonsimultaneously to the two channels, greatly reducing electric field interactions that might compromise perceived distinctions between channels (Wilson et al., 1987a; White et al., 1984).

Variations of the Breeuwer/Plomp processors are produced with different choices of parameters for the post processor. These parameters include (a) the duration of stimulus pulses for each channel; (b) the interval between pulses on different channels; (c) the order in which the channels are to be stimulated; (d) the mapping law for each channel, as described above; (e) the waveforms of stimulus pulses; and (f) whether stimulus cycles are to occur continually or are to be timed according to information provided by the voicing detector. Parameters a through c define the basic sequence of stimulation across the two channels, which we refer to as one update cycle. Update cycles can be repeated as rapidly as possible if voicing information is not to be explicitly coded, an approach constituting the most straightforward adaptation of the Breeuwer/Plomp strategy for cochlear implants. Alternatively, inputs from the voicing detector can be used to time the beginning of each update cycle. If voicing information is to be explicitly coded, cycles are timed to start in synchrony with the fundamental frequency (F0) during voiced speech sounds and at either randomly-varied or maximum-rate intervals during unvoiced speech sounds. Explicit coding of voicing information might be expected to improve a patient's perception of prosodic features associated with F0 contours and to help the patient make voice/unvoice distinctions for consonants (e.g. improve the ability to distinguish a "z" from an "s" or a "d" from a "t"). Also, an explicit representation of voicing information might be expected to improve the "naturalness" of speech percepts and the ability to make man/woman/child distinctions.

Fig. 2 shows typical waveforms for one variation of a Breeuwer/Plomp processor, as adapted for use in a cochlear prosthesis. In each panel the top trace is the input to the processor and the remaining traces are channel outputs. The input is the word "BOUGHT." The initial consonant occurs at about 180 msec 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" of the final consonant begins slightly before 640 msec, and an expanded display of waveforms beginning at 640 msec is shown in the lower-right panel. The lower panels thus exemplify differences in waveforms for voiced and unvoiced intervals.

In the particular variation of Breeuwer/Plomp processors presented in Fig. 2, balanced biphasic pulses are used, and voicing information is explicitly coded. During voiced speech sounds the update cycles are timed to begin in synchrony with the detected fundamental frequency, while during unvoiced speech segments they are initiated as rapidly as possible (i.e., at maximum-rate intervals). The pulse sequence in each cycle is such that the

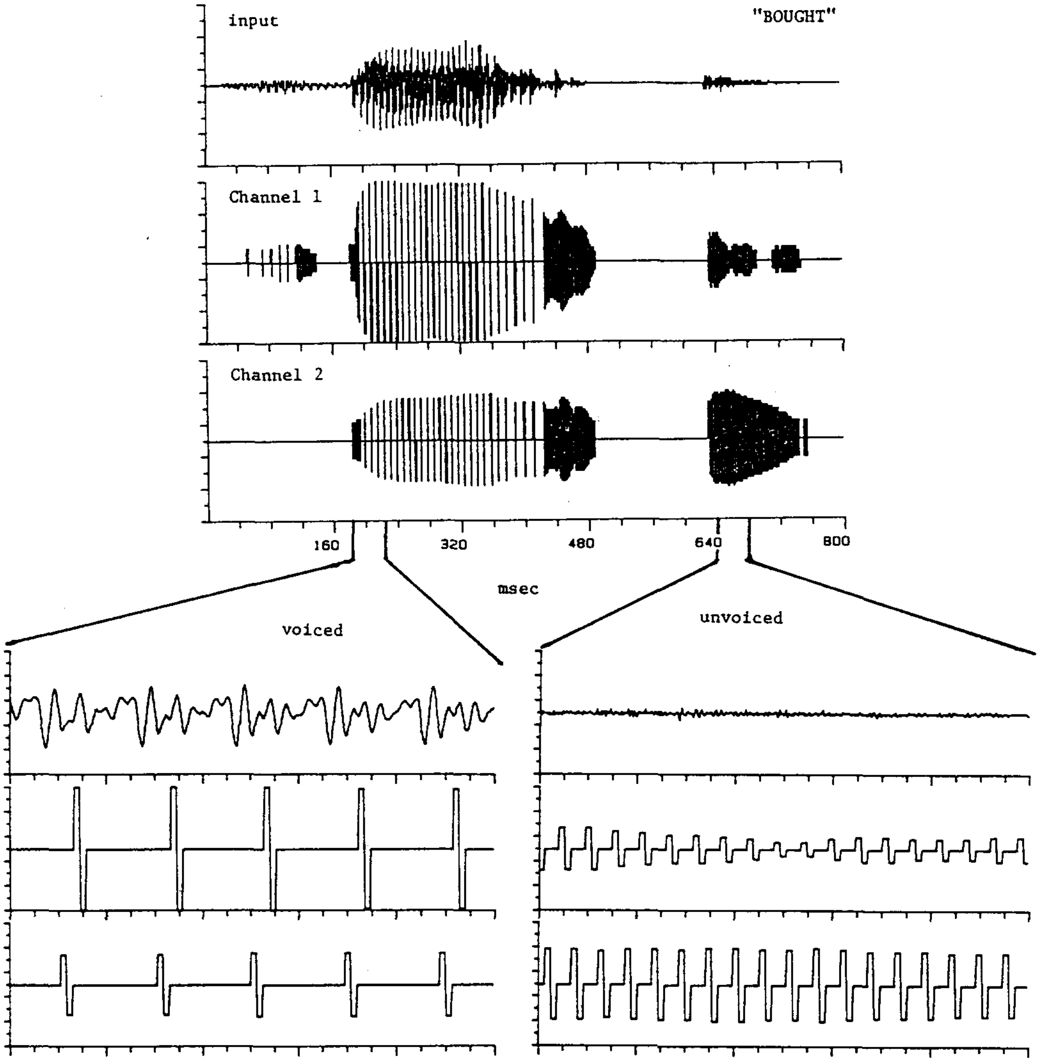


Fig. 2. Waveforms of a two channel, "Breeuwer/Plomp" processor. The top trace in each panel is the input to the processor and the remaining traces are channel outputs. The input is the word "BOUGHT." An expanded display of waveforms during the initial portion of the vowel is shown in the lower left panel and an expanded display of the waveforms during the "T" is shown in the lower right panel.

more basal electrode is stimulated first.

Tests and Procedures

Tests of vowel and consonant identification were conducted for all subjects. The tests were administered and interpreted in a "confusion matrix" format. The confusion matrix for vowels included the tokens "BOAT," "BEET," "BOUGHT," "BIT," and "BOOT," and the confusion matrix for consonants included the nonsense tokens "ATA," "ADA," "AKA," "ASA," "AZA," "ANA," "ALA," and "ATHA." The tokens in the vowel matrix were selected to measure the ability to perceive differences in the second formant (F2) of the vowels, and tokens in the consonant matrix were selected to measure the ability to distinguish the nonvisible consonants that have the greatest frequency of occurrence in spoken English (Schubert, 1985).

All implementations of Breeuwer/Plomp processors for the vowel and consonant identification tests were done with computer simulations as previously described (Wilson and Finley, 1985; Wilson et al., 1987a). The presentation of each processed token was accompanied by a display of response options on a computer video terminal used by the subject. Each response caused the updating of a matrix display on the investigator's computer screen (not seen by the subject), and the drawing of the next token from a randomized list. Five presentations of each processed token were included in the vowel test and three in the consonant test. At the end of a test the subject usually was given the overall correct score and an indication of the principal confusions made during the test. No feedback was given during the test itself. These short tests were oftentimes repeated after various intervals (e.g., on different days) to evaluate the reliability of the results. Because no significant differences were found among results across such repetitions, the responses were pooled for each subject and condition for which more than one test was conducted.

The test conditions included (a) vowel identification with lipreading; (b) vowel identification without lipreading; (c) consonant identification with lipreading; and (d) consonant identification without lipreading. Lipreading information was provided by miming tokens in synchrony with stimulus presentations. This information was provided by CCF for all tests with subject MH (the cable patient) and by DTL for all tests with the remaining subjects. Presentations of processed tokens usually were repeated at regular intervals until the subject responded. Although there is evidence that repetition of test tokens can increase scores (particularly for tests using open set material; see Merzenich et al., 1986 and Schindler et al., 1987), we did not find statistically-significant differences in the scores of several tests of consonant identification for single- and multiple-trial presentations.

In addition to the vowel and consonant tests just outlined, the Breeuwer/Plomp 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 input (De Filippo and Scott, 1978; Owens and Telleen, 1981; Owens and Raggio, 1987), and the IOWA test of medial consonant identification with video presentations of the speaker's face (Tyler et al., 1983). Because these tests were lengthy, and in some cases required direct interaction between investigator and subject (e.g., for the tracking test), a real-time implementation of the processor was used instead of the computer simulations. The design and use of the real-time processor has been described in previous reports (Finley et al., 1986 and 1987).

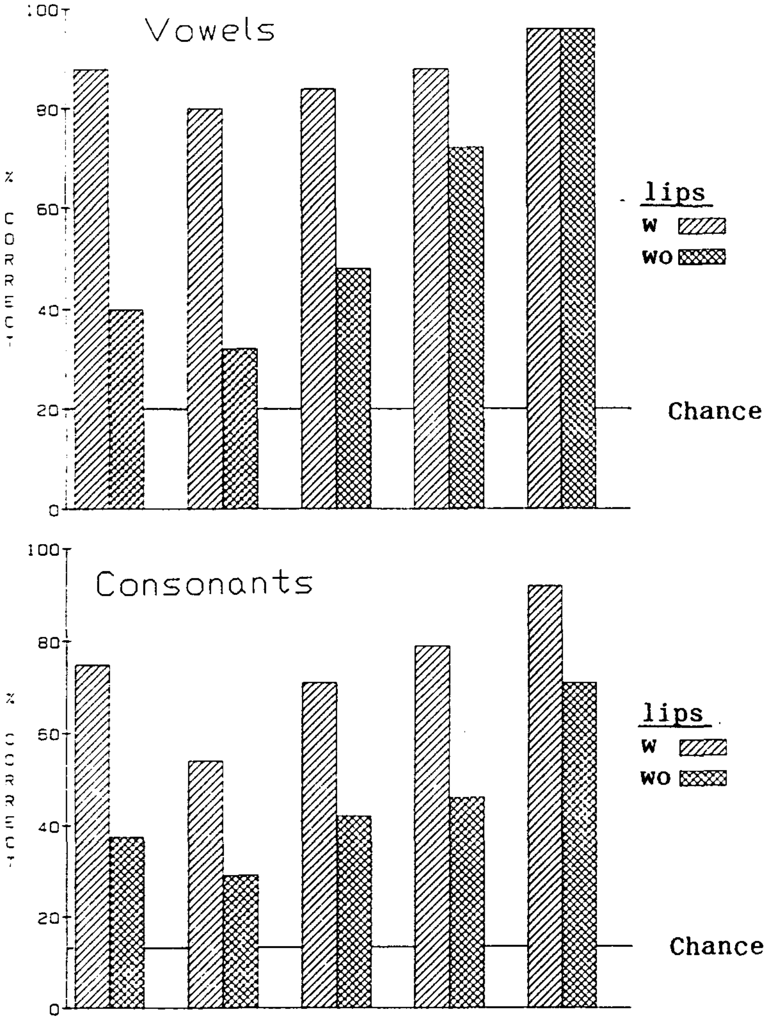
C. Results

Pilot Study

Our initial application of the Breeuwer/Plomp processor was in tests with subject MH (Wilson et al., 1986), who was fitted with a percutaneous cable. Nonsimultaneous pulsatile stimuli were delivered to the apical-most pair and the pair 4 mm basal to the apical-most pair in her implanted electrode array. Voicing information was not explicitly coded. Results from this pilot study are presented in Fig. 3, along with results from other processing strategies evaluated with MH. For each processor tested, at least 4 variations were evaluated to optimize processor parameters. The rationale and procedures for parametric manipulations have been presented in detail elsewhere (Wilson et al., 1986). The scores for each processor in Fig. 3 are those for the parametric set that produced the highest overall percent-correct score in the four tests of vowel and consonant identification.

Before describing the results for individual processors in Fig. 3, we note a few general features of the data. First, high scores are consistently found for the tests of vowel identification with lipreading. MH got 92% correct on a test we administered to measure her performance with lipreading alone, a score that is not significantly different from most of the scores shown in Fig. 3 for vowel identification using the combined input of lipreading plus auditory cues from a speech processor. Therefore, the scores for this vowel test with lipreading are not a sensitive indicator of processor performance.

Next, we note that scores on the tests of consonant identification using auditory and visual cues are a sensitive indicator of processor performance. In multiple tests of consonant identification with lipreading



Proc:	BP	CA	IP	IP	IP
Chans:	2	4	4	4	6
FO, v/uv?	N	--	N	Y	Y
Overall %	60	49	61	71	89

Fig. 3. Results of vowel and consonant identification tests for subject 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 BP for "Breeuwer/Plomp," CA for "compressed analog," and IP for "interleaved pulses"); the number of stimulation channels; whether voicing information was explicitly coded for the BP and 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.

only, MH got an average score of 52% correct. With the exceptions of the four channel, "compressed analog" (CA) processor all scores in Fig. 3 for consonant identification using lipreading plus processor are significantly above this level.

Third, test/retest reliability was good for MH. When we retested a processor that produced low scores on a previous occasion MH always would obtain low scores again, and when we retested a processor that produced high scores on a previous occasion MH 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 3%.

Returning now to comparisons of results across processors, Fig. 3 shows that the two-channel, Breeuwer/Plomp processor performs surprisingly well. All scores for vowel and consonant identification are significantly above chance, and the score for consonant identification using lipreading plus processor is significantly higher than the score for consonant identification with lipreading alone. Moreover, the overall scores for this processor are at least somewhat higher in every category when compared with the scores of the 4-channel CA processor. This CA processor is the one used in the UCSF/Storz clinical prosthesis. Thus, for patient MH the 2-channel Breeuwer/Plomp processor produces better results in tests of vowel and consonant identification than the clinically-applied four-channel processor. The largest increase is in the category of consonant identification with lipreading, as might be predicted from Breeuwer and Plomp's findings with normal-hearing subjects.

Additional comparisons among the results for the different processors presented in Fig. 3 indicate that scores are nearly identical for the 2-channel Breeuwer/Plomp processor and a 4-channel interleaved-pulses (IP) processor that represents the spectrum of speech in logarithmically-spaced bands between 350 and 3500 Hz (Wilson et al., 1986 and 1987a,b). However, if voicing information is explicitly coded in the 4-channel IP processor, the overall score jumps by 10%. This finding suggests that explicit coding of voicing information might also produce improved scores for the Breeuwer/Plomp processor. Finally, scores are further improved when the number of stimulation channels is increased from four to six. The best overall score presented in Fig. 3 is that for a 6-channel IP processor using explicit coding of voicing information.

The main conclusions from the pilot study conducted with MH are that (a) the two-channel Breeuwer/Plomp processor provides a surprising amount of useful information for this patient, especially for the condition of consonant identification using lipreading plus processor; (b) explicit coding of voicing information might improve the performance of the basic Breeuwer/Plomp processor; and (c) for processors without explicit coding of voicing information, the results from the tested variation of the Breeuwer/Plomp processor are superior or equivalent to the results obtained with all other tested processors using four or fewer channels of stimulation.

Vowel and Consonant Identification Tests

These encouraging results with patient MH provided the impetus to evaluate the Breeuwer/Plomp processor in tests with a variety of cochlear implant patients. The five subjects fitted with the 4-channel UCSF/Storz TTS participated in this follow-up study. As mentioned before, two channels with at least moderately-good isolation were selected for stimulation for each subject. The processors for all subjects used explicit coding of voicing information. Additional processor parameters for these subjects are shown in Table I.

TABLE I. Parameters selected for Breeuwer/Plomp processors.

subject	channels[a]	pulse widths/ phase[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.

Results from the tests of vowel and consonant identifications are presented in Fig. 4 in the form of confusion matrices. The patterns of responses from all five subjects are combined in the matrices for each condition; and the conditions include lipreading only, lipreading plus processor, and processor only. As with patient MH, the major improvements produced by the application of the Breeuwer/Plomp processor are in the categories of consonant identification. For example, the percentage of correct responses for consonant identification jumps from 41% correct to 81% correct when the auditory cues from the processor are added to the visual cues provided by lipreading. 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 shows that "l" and "th" are highly visible on the lips while the other consonants are not. The percentage of correctly identified consonants other than "l" and "th" is 22% for lipreading alone and 75% for lipreading plus processor. This huge improvement clearly demonstrates the potential utility of the Breeuwer/Plomp processor as a powerful adjunct to lipreading.

In addition, the Breeuwer/Plomp processor supports a high degree of vowel and consonant recognition with hearing alone. Specifically, the overall scores in these categories for the five subjects in this series 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 the hearing-only mode (Breeuwer and Plomp, 1984). The better results in our series might be attributable to the explicit coding of voicing information in our implementations of a modified "Breeuwer/Plomp" processor or to the relatively-small numbers of tokens used in the vowel and consonant tests. In either case it is noteworthy that the score for consonant identification with hearing alone is substantially higher than the score obtained with lipreading alone (62 vs. 41% correct),

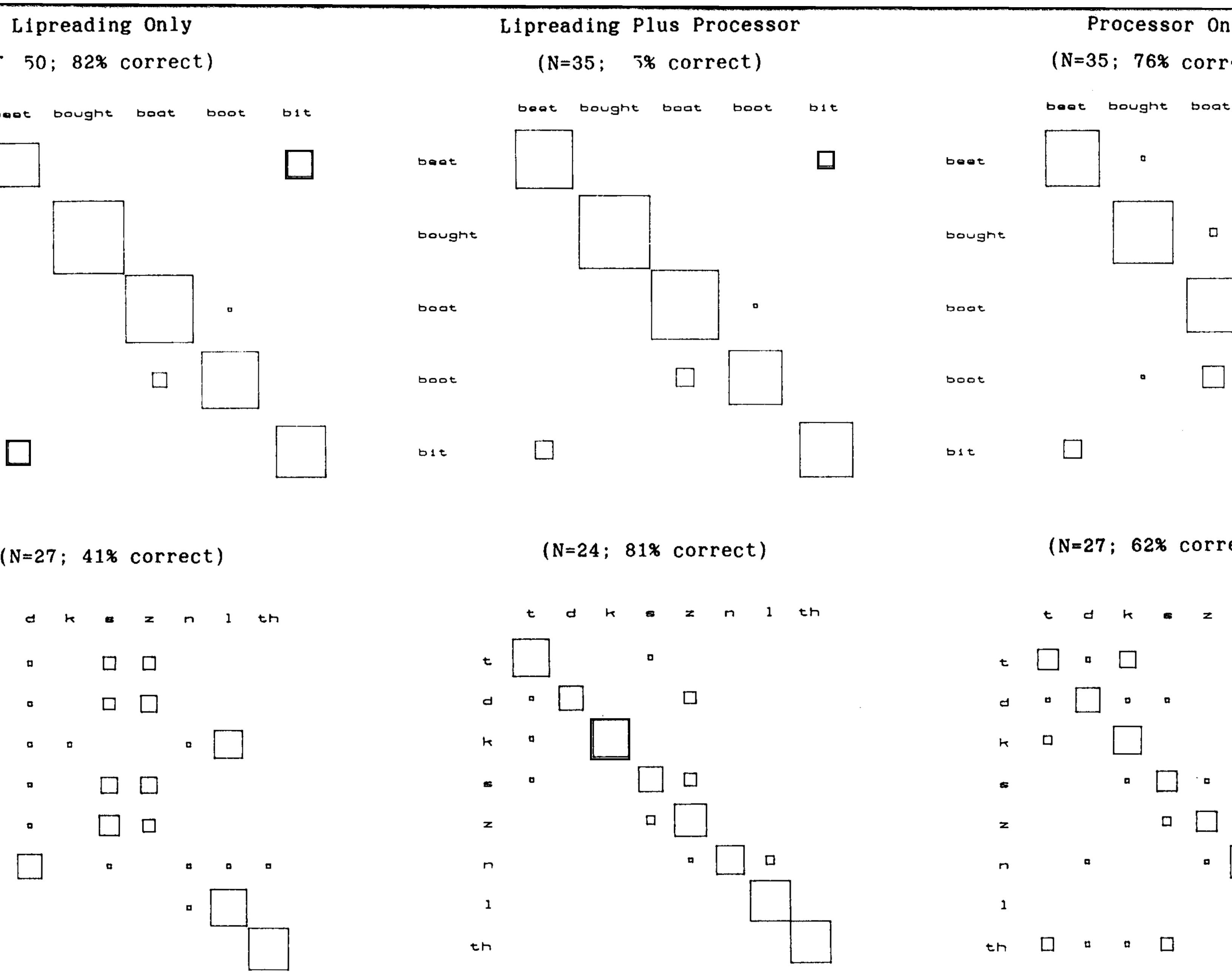


Fig. 4. Results of vowel and consonant identification tests for the five subjects fitted with the UCSF/Storz transcutaneous transmission system. The lengths of the sides of each square represent the averages of results obtained with two channel, Breeuwer/Plomp processors across all five subjects. Rows in the matrices represent stimuli and columns the responses. The total number of stimulus presentations and the percentage of correct responses are indicated in the parentheses above each matrix.

and that the score for vowel identification with hearing alone is only somewhat lower than the "topped out" score for lipreading alone (76 vs. 82% correct). Finally, as in the case of patient MH, the scores for vowel identification with lipreading and vowel identification using lipreading plus processor are indistinguishable.

Returning to our evaluation of the Breeuwer/Plomp processor as an adjunct to lipreading, Table II presents the scores for lipreading alone (L), lipreading plus processor (L+P), and processor alone (P) for all five subjects. In addition, the scores for the L+P and P conditions are compared to the scores for the L conditions in terms of percent change. A positive percent change thus indicates a benefit of the processor compared to the condition of lipreading alone. Finally, an aggregate score of "overall benefit" is calculated for each patient by averaging the percent change scores for the L+P and P conditions, and another aggregate score of "benefit to lipreading" is calculated by averaging the percent change scores for only the L+P conditions.

As might be expected from the average results across patients already presented in Fig. 4, large improvements are demonstrated in Table II for consonant identification when the auditory cues provided by the processor are added to the visual cues provided by lipreading (i.e., the L+P condition). These improvements range from 56% for subject HE to 129% for subject MC2. The average improvement across subjects is 101%. This two-fold improvement, along with the large improvements obtained for each of the individual subjects, suggests that applications of the Breeuwer/Plomp strategy can be expected to augment consonant identification for most cochlear implant patients.

The scores of percent change for the remaining conditions in Table II are, of course, lower than those just reviewed for consonant identification

TABLE II. Comparisons of Breeuwer/Plomp results with lipreading only data.[a,b]

	VOWELS[c]					CONSONANTS[d]					Overall	Benefit
subject	L	L+P	%	P	%	L	L+P	%	P	%	Benefit[e]	Lipread
MC1	80	76	- 5	70	-13	38	83	+118	58	+ 53	+38	+57
HE	74	64	-14	52	-30	48	75	+ 56	38	- 21	- 2	+21
RC	88	88	0	88	0	44	76	+ 73	67	+ 52	+31	+37
ET	82	96	+17	72	-12	44	100	+127	79	+ 80	+53	+72
MC2	88	100	+14	96	+ 8	31	71	+129	71	+129	+70	+72
Average Change =			+ 2		- 9			+101		+ 59		

[a] Abbreviations used in this table are: L for lipreading alone; L+P for lipreading plus processor; % percent change from lipreading alone score; and P for processor alone.

[b] All numeric entries are expressed in terms of percent correct or percent change.

[c] Chance performance is 20.0%.

[d] Chance performance is 12.5%.

[e] Overall benefit is defined as the average of all %'s for each subject.

[f] Benefit to lipreading is defined as the average of %'s for the two L+P conditions for each subject.

using lips plus processor. For example, the percent change scores for vowel identification using lips plus processor range from -14% for subject HE to +17% for subject ET. The average of the scores across subjects for this condition is +2%. These results indicate that small differences between the L and L+P conditions might be found for identification of vowels by individual subjects. As for subject MH, the vowel identification test lacks sensitivity because performance in the lipreading alone condition is so high. A more difficult test with a larger set of tokens might better reveal differences (if they exist) in the results obtained with the L and L+P conditions. The present data indicate that the high scores obtained in the lipreading alone condition are generally maintained with the addition of the processor.

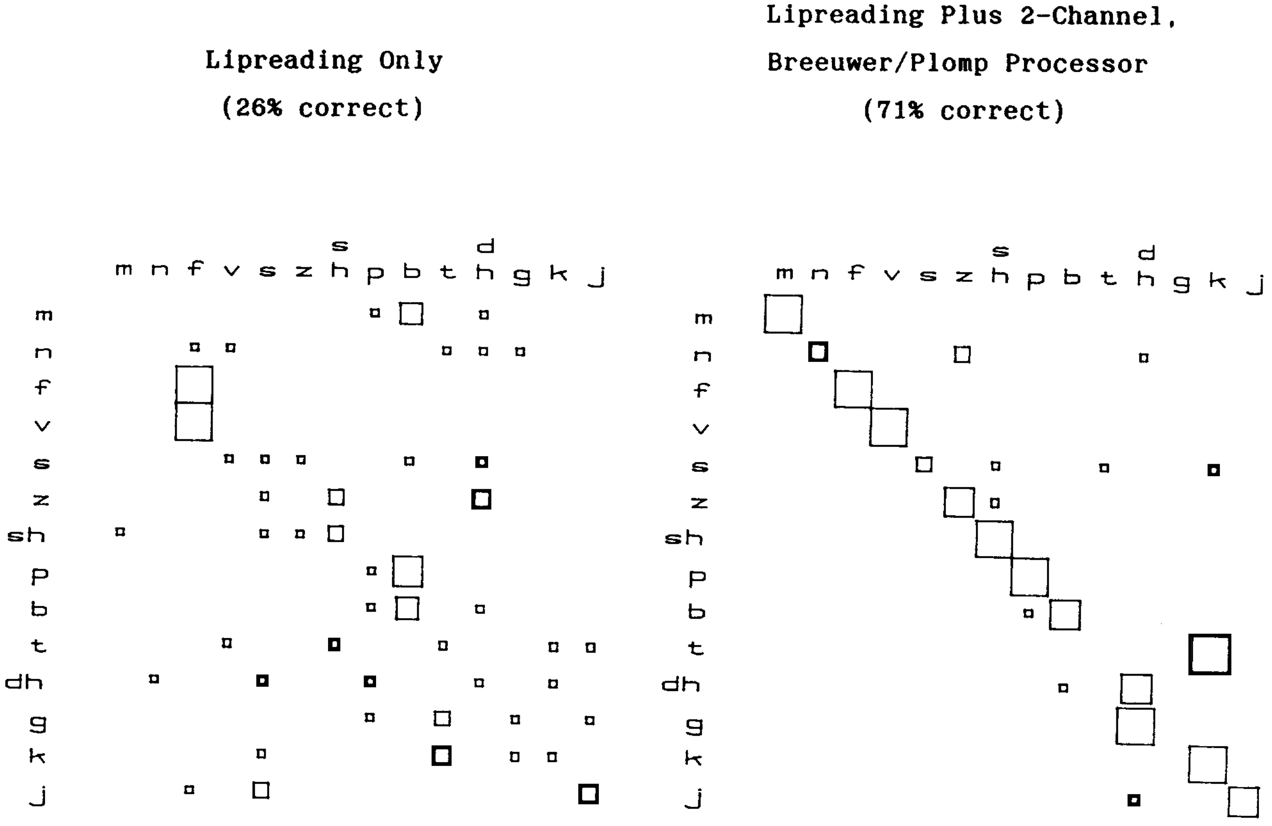
Because consonant identification is much more important than vowel identification for the perception of connected speech (see, e.g., Schubert, 1985), the actual "benefit to lipreading" provided by the Breeuwer/Plomp processor is probably greater than that indicated in the final column of Table II. In any case, the averages of percent change scores for the L+P conditions of vowel and consonant identification are positive for all subjects. These "benefit to lipreading" scores range from 21% for subject HE to 72% for subjects ET and MC2. The Breeuwer/Plomp processor thus produces substantial gains over lipreading alone for every tested subject.

The final sets of results in Table II are for vowel and consonant identification using the processor alone. For these conditions we see a modest decrease in most scores for vowel identification compared with lipreading alone (-9% overall, with a range of -30 to +8%), and a substantial increase in most scores for consonant identification compared with the scores for lipreading alone (+59% overall, with a range of -21 to +129%). In general, then, use of the processor in a hearing-only mode appears to provide somewhat more information than can be obtained by lipreading alone. All subjects except HE have combined scores for hearing only that exceed their combined scores for lipreading only. The overall benefit of the processor, as measured by averaging all the percent change scores for each subject, ranges from -2% for subject HE to +70% for subject MC2.

Additional Studies

Although the results just reviewed for the tests of vowel and consonant identification are most encouraging, especially for the important condition of consonant identification using the processor with lipreading, these tests sample a rather limited set of attributes associated with speech perception. We therefore decided to extend and confirm the results presented in Fig. 4 and Table II with a broad spectrum of speech perception tests, for which subject RC was available. Reference to Table II shows that RC ranked fourth among the five subjects in terms of overall benefit and in terms of benefit to lipreading. His results from the vowel and consonant identification tests are thus somewhat below the averages of results obtained for all subjects in this series.

As indicated above, the additional tests conducted with RC included all subtests of the MAC battery, connected discourse tracking with and without the processor, and the IOWA test of medial consonant identification with lipreading cues. The results from the IOWA test are presented in Fig. 5. As with the previous consonant identification tests, the consonants are presented in an a<u>C</u>a context (e.g., "AFA"). Unlike the previous tests, though, the lipreading cues are completely controlled in the IOWA tests in that these cues are presented from a videotape recording of the speaker's face. Also, many more consonants are included in the IOWA test (i.e., 14 vs. 8).



Lipreading Only
(26% correct)

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

Fig. 5. Results from IOWA tests of consonant identification with visual 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 two-channel Breeuwer/Plomp processor. Rows in the matrices represent stimuli and columns the responses. Each token was repeated five times for each test.

The results from the IOWA test (Fig. 5) more than confirm the previous results from the 8-token identification test. RC obtains a score of 26% correct for lipreading alone on the IOWA test and a score of 71% correct for lipreading plus processor, an almost 3-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 a modest amount of training or learning. In particular, "t" was always perceived as "k" but never vice versa, and "g" was always perceived as "dh" but never vice versa. These asymmetries in the errors indicate that information is available to make the distinctions, but is not being used by the subject. Training can made subjects aware of the information and teach them how to utilize it in making the appropriate distinctions. Elimination of the t/k and g/dh errors in RC's matrix for the lipreading plus processor condition would boost his score to 86% correct. Even without this increase, though, the results in Fig. 5 demonstrate that the great majority of frequently-occurring consonants in English are well represented by the combined input of lipreading plus speech processor.

Results from the tests of connected discourse tracking further extend and 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 an extremely poor level of performance to a moderately good level of performance. His tracking rate with the processor is about half the average rate obtained in tests with subjects who have normal hearing (Owens and Raggio, 1987). The use of the Breeuwer/Plomp processor would thus allow RC to understand speech at acceptable rates, especially for typical conversations with a high content of contextual information.

The remaining tests conducted with RC were those of the full MAC battery. The results, presented in Table III, indicate substantial access to speech information in the hearing-only mode. All scores from the closed set tests of prosodic perception and of phoneme and word discrimination are significantly above chance at the P = .05 level. 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 are also obtained for the Accent (80% correct), Vowel (62% correct) and Final Consonant (63% correct) tests. These high scores are unexpected inasmuch as the Breeuwer/Plomp 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 Breeuwer/Plomp processor can provide useful information even when the visual cues from lipreading are not available. In particular, very 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._, 1987). We note, however, that scores on open-set tests are likely to reflect the cognitive and linguistic skills 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 Breeuwer/Plomp processor in the hearing-only mode. That is, the information provided by this processor is certainly adequate to support very high levels of open-set recognition; however it is unlikely that all implant patients will be able to utilize this information as effectively as RC.

TABLE III. Results from the Minimal Auditory Capabilities (MAC) battery for subject RC.

Tests	Chance (%)	Score (%)[a]
PROSODIC (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 PROSODIC and PHONEME & WORD DISCRIMINATION tests are significantly above chance at the P=.05 level; any score above zero is regarded as significant for the OPEN-SET SPEECH RECOGNITION tests.

D. Discussion

The results of this study confirm and extend the findings of Breeuwer and Plomp. In particular, the present results from our evaluation of "Breeuwer/Plomp" processors for cochlear implants demonstrate the following: (a) the Breeuwer/Plomp processor can act as a powerful adjunct to lipreading, as evidenced by large improvements in consonant identification (six subjects) and connected discourse tracking (the one subject tested) when the processor is used in conjunction with lipreading; (b) this aid in consonant identification is likely to apply to most implant patients, as indicated by the improvements obtained for every tested subject in our series of six; (c) in addition to its utility as an adjunct to lipreading, the Breeuwer/Plomp processor supports a high degree of vowel and consonant identification with hearing alone; and (d) the high scores of the one subject tested with the MAC battery further indicate that the Breeuwer/Plomp processor can provide useful information even when visual cues from lipreading are not available. These promising results should encourage consideration of the use of a Breeuwer/Plomp processor in cochlear implants whenever the number of channels available for distinct stimulation is restricted.

E. Acknowledgments

We thank the patients who participated in this study for their dedicated effort and pioneering spirit. We are also pleased to acknowledge the important contributions of M.W. Skinner, who collaborated with us in the pilot studies with patient MH; D.K. Kessler, who made the arrangements for visits to Duke by patients MC1, HE, RC, ET and MC2; and R.D. Wolford, who conducted the MAC and speech-tracking tests with patient RC. In addition to the contributions of these individuals, we are most grateful for the assistance and informed guidance provided by M.W. White, M.M. Merzenich, B.A. Weber, J.C. Farmer, Jr., R.A. Schindler, G.E. Loeb, F.T. Hambrecht and P.D. Kenan.

F. References

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

De Filippo, C.L., and Scott, B.L. (1978). "A method for training and evaluating the reception of ongoing speech," J. Acoust. Soc. Am. 63, 1186-1192.

Finley, C.C., Wilson, B.S., and Lawson, D.T. (1986). "Speech processors for auditory prostheses," Third Quarterly Progress Report, NIH project N01-NS-5-2396 (National Institutes of Health, Bethesda, MD).

Finley, C.C., Wilson, B.S., and Lawson, D.T. (1987). "Speech processors for auditory prostheses," Seventh Quarterly Progress Report, NIH project N01-NS-5-2396 (National Institutes of Health, Bethesda, MD).

Gantz, B.J., McCabe, B.F., Tyler, R.S., and Preece, J.P. (1987). "Evaluation of four cochlear implant designs," Ann. Otol. Rhinol. Laryngol. 96 (Suppl. 128), 145-147.

Loeb, G.E., Byers, C.L., Rebscher, S.J., Casey, D.E., Fong, M.M., Schindler, R.A., Gray, R.F., and Merzenich, M.M. (1983). "Design and fabrication of an experimental cochlear prosthesis," Med. & Biol. Eng. & Comput. 21, 241-254.

Merzenich, M.M., and White, M.W. (1977). "Cochlear implant: The interface problem," in Functional Electrical Stimulation: Applications in Neural Prostheses, edited by F.T. Hambrecht and J.B. Reswick (Marcel Dekker, New York), pp. 321-346.

Merzenich, M.M., Kessler, D.K., Yanda, J.L., and Rebscher, S.J. (1986). "Widespread use of nonconventional audiological testing procedures for evaluation of cochlear implant patients," Proc. Assoc. Res. Otolaryngol. 9th Midwinter Res. Conf., p. 70.

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

Owens, E., and 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.

Owens, E., and Telleen, C. (1981). "Tracking as an aural rehabilitative process," J. Acad. Rehabil. Audiol. 14, 259-273.

Schindler, R.A., Kessler, D.K., Rebscher, S.J., Jackler, R.K., and Merzenich, M.M. (1987). "Surgical considerations and hearing results with the UCSF/Storz cochlear implant," Laryngoscope 97, 50-56.

Schindler, R.A., Kessler, D.K., Rebscher, S.J., Yanda, J.L., and Jackler, R.K. (1986). "The UCSF/Storz multichannel cochlear implant: Patient results," Laryngoscope 96, 597-603.

Schubert, E.D. (1985). "Some limitations on speech coding for implants," in Cochlear Implants, edited by R.A. Schindler and M.M. Merzenich (Raven Press, New York), pp. 269-276.

Tyler, R.S., Preece, J.P., and Lowder, M.W. (1983). "The Iowa cochlear-implant test battery," Laboratory Report (Department of Otolaryngology -- Head and Neck Surgery, University of Iowa at Iowa City).

van den Honert, C., and Stypulkowski, P.H. (1987). "Single fiber mapping of spatial excitation patterns in the electrically stimulated auditory nerve," Hear. Res. 29, 195-206.

White, M.W., Merzenich, M.M., and Gardi, J.N. (1984). "Multichannel cochlear implants: Channel interactions and processor design," Arch. Otolaryngol. 110, 493-501.

Wilson, B.S., and Finley, C.C. (1985). "A computer-based simulator of speech processors for auditory prostheses," Proc. Assoc. Res. Otolaryngol. 8th Midwinter Res. Conf., p. 109.

Wilson, B.S., Finley, C.C., and Lawson, D.T. (1986). "Speech processors for auditory prostheses," Second Quarterly Progress Report, NIH project N01-NS-5-2396 (National Institutes of Health, Bethesda, MD).

Wilson, B.S., Finley, C.C., Farmer, J.C., Jr., Lawson, D.T., Weber, B.A., Wolford, R.D., Kenan, P.D., White, M.W., Merzenich, M.M., and Schindler, R.A. (1987a). "Comparative studies of speech processing strategies for cochlear implants," Laryngoscope (accepted for publication).

Wilson, B.S., Finley, C.C., Lawson, D.T., and Wolford, R.D. (1987b). "Speech processors for cochlear prostheses," submitted for publication.

III. Plans for the Next Quarter

Our plans for the next quarter include the following:

1. Continue ongoing psychophysical and speech perception studies with patient MH, with an emphasis on collaborative work with Bryan Pfingst of the Kresge Hearing Research Institute to extend our previous investigations of pitch and loudness coding with cochlear implants;
2. Continue our collaboration with UCSF to develop the speech processor and transcutaneous transmission system for a next-generation auditory prosthesis;
3. Begin collaborative studies with Sigfrid Soli and others at the 3M company, to evaluate alternative processing strategies for two patients implanted with a scala-tympani array of four pairs of longitudinally-oriented bipolar electrodes (the first patient will be studied at Duke during the period of Nov. 30 to Dec. 9, 1987); and
4. Present project results in (a) one talk at the *Neural Prosthesis Workshop*, Oct. 28-30, and (b) four talks at the *Ninth Annual Conference of the IEEE Engineering in Medicine and Biology Society*, Nov. 13-16.

Appendix 1

Summary of Reporting Activity for the Period of

June 27 through September 26, 1987, NIH Contract N01-NS-5-2396

The following presentations were made in the present reporting period:

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.

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.

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

Wilson, B.S.: Chairman of speech coding panel, International Cochlear Implant Symposium 1987, Düren, West Germany, Sept. 7-11, 1987.

The published abstract for the fourth presentation (by Schindler and Wilson) is reproduced on the next page of this appendix.