Original sent to printing

Fourth Quarterly Progress Report

February 1 through April 30, 1990

NIH project N01-DC-9-2401

"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 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. Work in the present quarter included the following:

1. Initiation of studies with a series of eight Symbion patients, to evaluate the continuous interleaved sampler (CIS), peak-picker and other processors across a population of subjects fitted with the Symbion electrode array and percutaneous connector.
2. Studies with two Nucleus patients, to evaluate reduced implementations of the CIS processor with subjects fitted with the Nucleus electrode array and transcutaneous transmission system (TTS).
3. Studies with one UCSF/Storz patient, to evaluate reduced implementations of the CIS processor with a subject fitted with the UCSF/Storz electrode array and TTS.
4. Presentation of project results at the *13th Midwinter Research Conference of the Association for Research in Otolaryngology* (Finley, Feb. 4-8), the annual *AAAS Meeting* in New Orleans (Wilson, Feb. 20), and at Richards Medical Company in Memphis (Wilson, Feb. 5).
5. Continued preparation of manuscripts for publication.

In this report we present results to date on the comparison of the compressed analog (CA) and CIS processors in tests with subjects fitted with the Symbion device (point 1 above). These results demonstrate superiority of the CIS processor for all seven of the subjects studied thus far. Indeed, the CIS processor provides superior scores on every administered test of open-set recognition (spondees, CID sentences, SPIN sentences, NU-6 words, and connected discourse tracking) for every subject, despite the multiple years of experience each subject had had with his or her CA processor. Results from the remaining subjects and processors in this series, along with the results from activities 2 and 3 above, will be presented in future reports.

II. Comparison of *Compressed Analog* and *Continuous Interleaved Sampler* Processors in Tests with Symbion Subjects

In intensive studies with subject MP (SR2) we evaluated four variations of interleaved pulses (IP) processors, four variations of continuous interleaved sampler (CIS) processors, and one variation of a peak picker (PP) processor. As described in QPR 2 for this project, one variation of the IP processor and one variation of the CIS processor (referred to as the "supersampler" processor in QPR 2) were compared with each other and with MP's compressed analog (CA) processor using a full battery of speech tests. Results from those comparisons indicated superiority of the CIS processor for MP. Indeed, unprecedented levels of speech recognition with a cochlear implant were obtained with that processor and subject.

To evaluate the generality of the results with MP, we initiated a series studies with seven additional subjects implanted with the Symbion device. The purpose of this report is to provide interim results, on comparison of the CA and CIS processors, for the first seven subjects (including MP) in the series.

Methods

Processors

Waveforms of the CA and CIS processors are shown in Fig. 1. Briefly, the CA processor first compresses the wide dynamic range of input speech signals into the narrow dynamic range available for electrical stimulation of the auditory nerve. The compressed signal then is filtered into frequency bands for presentation to each electrode. As can be appreciated from Fig. 1, CA stimuli contain many temporal details of the input speech signals. For example, in the left column strong periodicities in the apical two channels reflect the fundamental frequency (F0) and first and second formant frequencies (F1 and F2) of the voiced speech sound. The onset of an unvoiced /t/ burst is represented in the stimuli of the basal channels, as may be seen in the right column.

One concern associated with the use of CA processors is that of channel interactions [White et al., 1984; Wilson et al., 1988b]. Simultaneous stimulation of two or more channels with continuous waveforms results in summation of the electrical fields from the different electrodes. This summation can exacerbate interactions among channels, and thus may reduce the salience of channel-related cues.

Another concern is that many of the temporal details present in CA stimuli may not be perceived by implant patients. Most patients cannot perceive changes in the frequency of stimulation above a "pitch saturation limit" of about 300 Hz [e.g., Shannon, 1983]. Thus, while most patients may be able to perceive changes in F0, only exceptional patients will be able to make use of the F1 information contained in the stimuli for apical channels. It is highly unlikely that any patient would be able to perceive changes in F2 through temporal cues alone.

The problem of channel interactions is addressed in the CIS processor through the use of interleaved nonsimultaneous stimuli. There is no temporal overlap between stimulus pulses, so that

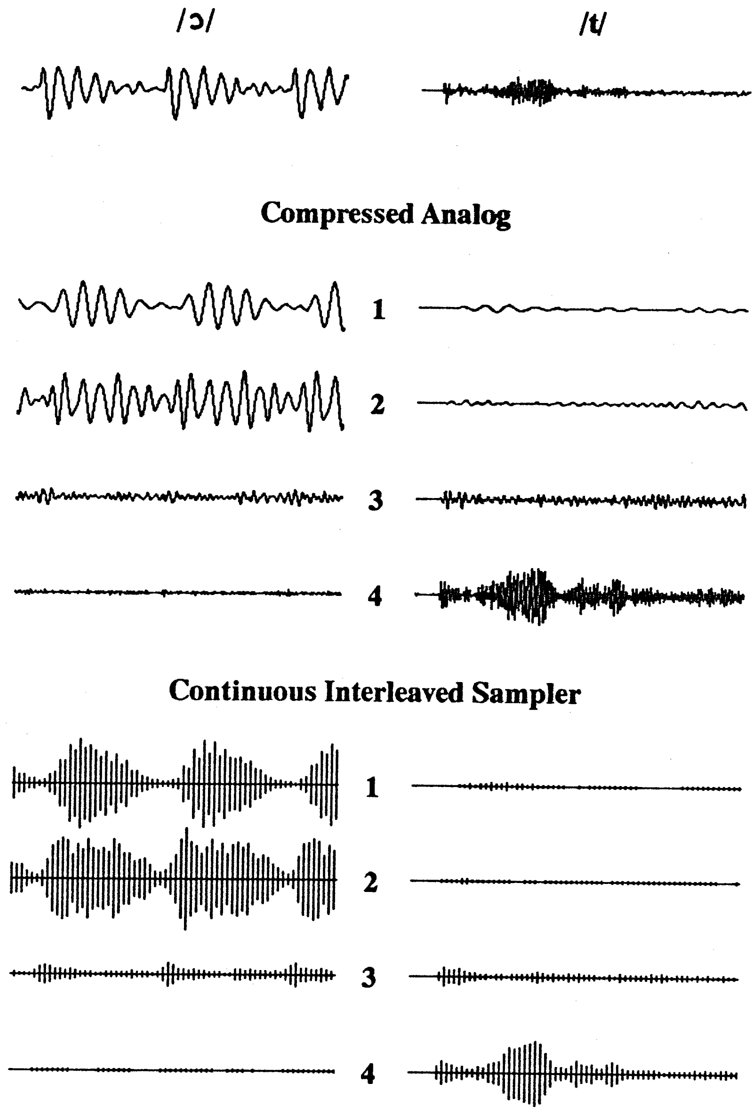

Fig. 1. Waveforms of the CA and CIS processors. Equalized (6 dB/octave attenuation below 1200 Hz) speech inputs are shown at the top and stimulus waveforms for each of the processors below. The left column shows input and stimulus waveforms for a voiced speech sound and the right column those for an unvoiced speech sound. Stimulus waveforms are numbered by channel, with channel 1 delivering its output to the apical-most electrode in the scala tympani. Center frequencies for the bandpass filters associated with channels 1-4 are 0.5, 1.0, 2.0, and 4.0 kHz respectively. The time constant of the integrating filters for bandpass energy detection in the CIS processor is 0.4 ms. The duration of each trace is 25.4 ms.

direct summation of electrical fields is avoided. The energy in each frequency band of the input signal is represented by the amplitudes of the pulses delivered to the corresponding electrode. The pulses shown in Fig. 1 have a one-to-one correspondence with the root-mean-square (RMS) energies in each band. In actual applications of the CIS processor, pulse amplitudes are determined with a logarithmic or power-law transformation of RMS energies to compress the dynamic range of those energies into the range of electrically-evoked hearing.

In contrast to the IP processor described in previous reports [Wilson et al., 1988a; 1988b; 1990], the CIS processor presents stimulation cycles at a constant, rapid rate during both voiced and unvoiced segments. In addition, this processor generally uses brief pulses, with one presented immediately after its predecessor, so that rapid variations in RMS energies can be followed by variations in pulse amplitudes for each channel. Some patients may be able to make use of this information to perceive changes in F1 and to perceive the rapid temporal variations important for the identification of certain consonants (variations up to about 200 Hz, see Van Tassell et al., 1987).

Subjects

Subjects for this series were selected on the basis of performance with the Symbion device. In particular, we identified (with the help of others, see Acknowledgements) a population of subjects with levels of performance similar to MP's. As indicated in the Results section, all subjects had scores of 30% or higher on the NU-6 test of monosyllabic word recognition with their CA processors. Such high scores are rare among implant patients.

Each subject was studied for a one-week period in which (a) basic psychophysical measures ~~with~~ ere obtained on thresholds and dynamic ranges for pulsatile stimuli, (b) a variety of CIS processors (with different choices of processor parameters) were evaluated with tests of consonant and vowel identification, and (c) performance with one or two CIS processors and the clinical CA processor was documented with additional tests.

Tests

The CA and CIS processors were evaluated with a variety of speech perception tests. Because the subjects had excellent performance with both strategies, only results from the most difficult sound-alone tests are reported here. These included identification of 16 consonants (/b, d, f, g, dʒ, k, l, m, n, p, s, ʃ, t, ð, v, z/) in an /a/-consonant-/a/ context; identification of 8 vowels (/i, ɔ, ɛ, u, I, U, ʌ, æ/) in a /h/-vowel-/d/ context; the segmental and open-set tests of the Minimal Auditory Capabilities (MAC) battery [Owens et al., 1985]; and connected discourse tracking [De Filippo and Scott, 1978; Owens and Raggio, 1987].

In both the consonant and vowel tests multiple exemplars of the tokens were played from laser videodisc recordings of male and female speakers [Tyler et al., 1987; Lawson et al., 1989]. A single block of trials consisted for five randomized presentations of each consonant or three randomized presentations of each vowel for one of the speakers. At least two blocks were administered for each speaker, processor and subject in the consonant tests, and at least three blocks were administered for each speaker, processor and subject in the vowel tests.

The segmental tests included identification of the word containing the correct vowel, initial consonant (Init Cons), or final consonant (Fnl Cons) among four options for each test item. The vowel test contained 60 items, the initial consonant test 64 items, and the final consonant test 52 items.

The open-set tests included recognition of 50 one-syllable words from Northwestern University Auditory Test 6 (NU-6); 25 two-syllable words (spondees); 100 key words in the Central Institute for the Deaf (CID) sentences of everyday speech; and the final word in 50 sentences from the Speech Perception in Noise (SPIN) test. In both the segmental and open-set tests single presentations of the words or sentences were played from cassette tape recordings of a male speaker.

In the tracking test the subject's task was to repeat verbatim previously-unknown paragraphs read by a trained speaker [Owens and Raggio, 1987]. For items not understood after the first presentation, various strategies such as repetition of phrases or words were used until the items were correctly repeated.

All tests were conducted with hearing alone, and all tests except tracking used single presentations of recorded material with no feedback as to correct or incorrect responses. The score for the tracking test was calculated by dividing the number of words in four paragraphs by the time taken to complete those paragraphs. Scores for the remaining tests were calculated as the percentage of correct responses. In addition, results for the consonant identification test were expressed as percent information transfer for articulatory and acoustic features that characterize the selected consonants [Miller and Nicely, 1955; Wilson et al., 1990].

Processor Parameters

Each subject's own clinical device (the Symbion prosthesis) was used for the tests with the CA processor. Four channels of stimulation were used for all subjects. Detailed descriptions of the clinical CA processor may be found in papers by Eddington [1980; 1983].

Selection of parameters for the CIS processors was guided by results from preliminary tests of consonant identification, primarily with the male speaker. The final choices for each subject are presented in Table 1. The processors for all subjects used pulses with durations of 102 μs or less, 5 or 6 channels of stimulation, and rates of stimulation above 800 Hz on each channel. In addition, the order of channels in the stimulation cycle was chosen to maximize the spatial separation between sequentially stimulated channels. We expected that this "staggered order" might produce a further reduction in channel interactions.

As indicated in Table 1, one CIS processor was evaluated with a full battery of speech tests for subjects SR2-6 and SR7, and two processors were evaluated for subjects SR5 and SR8. With the exception of the tracking test for subject SR4, all tests were conducted with the first processors listed for each subject. Tests with the second processor for subject SR5 were limited to the segmental and open-set tests of the MAC battery, and tests with the second processor for subject SR8 were limited to those tests and tracking.

Table 1. Parameters of CIS processors. The parameters include pulse duration per phase (μs/ph), the type of rectifier (Half Wave or Full Wave) used in the circuits for bandpass energy detection (RMS rect), the corner frequency of the integrating filters in those circuits (RMS filters), the frequency below which speech signals are attenuated for input equalization (eq), the sequence of channels for each stimulation cycle (channel sequence; channel 6 is the most basal for all subjects except subject SR5, see footnote a), the rate of pulsatile stimulation on each channel (rate), and the type of transformation used to map pulse amplitudes (mapping). The logarithmic transformation for mapping is of the form $pulse\ amplitude = A\ x\ log(RMS) + k$, and the power-law transformation is of the form $pulse\ amplitude = A\ x\ (RMS)^p + k$, where A and k are set so that pulsatile stimuli derived from processed speech will span the dynamic range from threshold to comfortable loudness on each channel.

Subject	μs/ph	RMS rect	RMS filters (Hz)	eq (Hz)	Channel sequence	rate (Hz)	mapping
SR2	55	FW	800	600	6-3-5-2-4-1	1515	logarithmic
SR3	31	FW	400	1200	6-3-5-2-4-1	2688	logarithmic
SR4	63	FW	400	1200	6-3-5-2-4-1	1323	logarithmic
SR5	31	HW	800	1200	2-5-4-6-1[a]	3226	logarithmic
	31	HW	800	1200	2-5-3-1-6-4[b]	2688	logarithmic
SR6	102	FW	400	1200	6-3-5-2-4-1	817	logarithmic
SR7	34	HW	400	1200	5-3-1-4-2[c]	2941	power law ($p = 0.2$)
SR8	100	FW	400	1200	6-3-5-2-4-1	833	logarithmic
	100	FW	400	1200	6-3-5-2-4-1	833	power law ($p = 0.2$)

[a]The electrodes for subject SR5 were inserted into the scala tympani one at a time, instead of as a bundled array. Because of uncertainties in the depths of insertion for the individual electrodes, the electrode positions had to be inferred on the basis of tonotopic ranking. The channel sequence from these inferred positions was 5-3-1-4-2. Electrode 3 was not used in this five-channel processor because stimulation of that electrode produced markedly different pitches at different stimulus levels.

[b]The channel sequence from the inferred positions of the electrodes was 6-3-5-2-4-1.

[c]Subject SR7 used a five-channel processor because stimulation of his sixth channel produced transient sensations of head movements.

Evaluation of Practice and Learning Effects

Because the initial tests with the CA processor preceded those for the CIS processor, we were concerned that practice or learning effects might favor the CIS processor in comparisons of the two strategies. To evaluate this possibility, tests of consonant identification with the CA processor were repeated at the end of the week for each subject. In all cases except one (subject SR3), the retest scores were indistinguishable from the original scores, and data from the second tests were added to those of the first. In the exceptional case, the retest scores were about 10% higher than the original scores for the male speaker and indistinguishable for the female speaker. The retest data for the male speaker, and combined test/retest data for the female speaker, were used in all subsequent analyses for subject SR3.

In addition, the CID sentence and NU-6 word tests were repeated with the CIS processor for five of the subjects (subjects SR3-7 and SR8) using a different recorded speaker and new lists of words and sentences. Practice or learning effects would be demonstrated by significant differences in the test/retest scores. However, no such differences were found ($p > 0.6$ for paired t comparisons of the CID scores; $p > 0.2$ for the NU-6 scores), and the scores from the first and second tests were averaged for all subsequent analyses.

Results

Results Across Subjects

Results across subjects are presented in Fig. 2 and in Tables 2 and 3. These results were obtained from the tests with the first processors listed in Table 1 for each subject. The raw scores from those tests are presented in Appendix Tables A.1 and A.2 for reference.

Table 2 and the first halves of Table 3 and Fig. 2 show the results from the tests of consonant and vowel identification. As mentioned before, the patterns of confusions and correct responses were evaluated using information transmission (IT) analysis. In this analysis the "relative transinformation" is calculated for selected articulatory or acoustic features of the phonemes in the 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 (dur), frication (fric), and envelope cues (envel). The vowel features were first formant frequency (F1), second formant frequency (F2), and duration (dur). The assignments of these features for the phonemes in the identification tests are presented in Appendix Tables A.3 and A.4.

Matrices of stimuli and responses were compiled for each subject and processor by combining the data for the male and female speakers. The combined matrices had a minimum of 20 trials for each of the consonants and 18 trials for each of the vowels.

The overall percent-correct scores and total number of trials for the tests of consonant and vowel identification are presented in Table 2. Paired-t comparisons of the scores demonstrate

Table 2. Results from the tests of consonant and vowel identification. Shown are the number of trials for each token (trials), the mean percent correct scores (% correct), and significance levels from paired t comparisons (p).

Test	Processor	Trials	% Correct	p
Consonant	CA	205	66.4	
	CIS	145	81.9	.01
Vowel	CA	132	95.1	
	CIS	126	92.7	NS

superiority of the CIS processor for consonant identification ($p < .01$) and no difference between processors for vowel identification.

Although percent correct scores provide a rough indication of processor performance, IT analyses can demonstrate specific strengths and weaknesses of a given strategy. Means of the scores from those analyses are presented in Table 3, along with the results from paired-t and Wilcoxon (nonparametric) comparisons of the scores for the two processors. IT scores for all consonant features except voicing are significantly higher with CIS processor, with especially large differences found for overall transmission, nasality, frication, and place of articulation. The IT scores for the vowel features are indistinguishable except for overall transmission, where the scores for overall transmission are higher for the CA processor ($p < .05$).

In addition to the IT analyses of matrices for each of the subjects, analyses were performed using aggregate matrices across subjects. These were compiled for each processor by combining the subject matrices, and are presented for reference in Appendix Tables A.5 and A.6.

Results from IT analyses of the aggregate matrices are shown in Fig. 2. Large gains in the consonant features of overall transmission, nasality, frication, and place of articulation again are seen when the CIS processor is used instead of the CA processor. In addition, substantial increases are found for consonant duration and envelope cues. Finally, notice that the absolute scores for most features approximate the ceiling of perfect performance when the CIS processor is used. The scores for nasality and envelope each exceed 80%, and the scores for all remaining features except place (67.3%) exceed 70%. The greatest strengths of the CIS processor are in the transmission of nasality, frication and envelope information, while the greatest strengths of the CA processor are in the transmission of voicing, duration and envelope information. A relative weakness shared by both processors is in the transmission of place information. Further weaknesses of the CA processor lie in the transmission of nasality and frication information.

Scores from IT analyses of the aggregate matrices for vowels approximate the ceilings of perfect performance for both processors and all features. Transmission scores are nearly identical

Table 3. Results of speech tests for the seven subjects. Means and standard deviations (SD) are shown for the CA and CIS processors. Levels of significance from paired t tests and Wilcoxon signed ranks (nonparametric) tests are indicated in the right columns.

	CA		CIS			
	Mean	SD	Mean	SD	Paired t	Wilcoxon
Consonants						
Overall	67.8	9.8	83.1	11.1	.01	.05
Voice	70.6	10.4	80.4	17.4		
Nasal	60.4	26.0	86.4	18.1	.01	.02
Fric	51.5	19.3	81.2	16.8	.01	.05
Dur	60.5	9.1	76.1	19.8	.05	
Place	55.7	8.5	72.9	19.3	.02	.05
Envel	71.7	14.6	86.8	12.1	.05	.05
Vowels						
Overall	93.0	8.0	88.9	5.9	.05	
F1	88.3	16.6	85.4	13.6		
F2	88.1	16.9	78.8	10.6		
Dur	90.6	13.4	91.5	6.2		
Segmentals						
Vowel	78.3	6.2	80.3	7.8		
Init Cons	92.7	5.5	94.6	3.5		
Fnl Cons	85.7	6.7	95.6	3.3	.02	.05
Open Set						
Spondee	76.0	16.2	94.3	8.6	.05	.05
CID	86.9	13.8	98.8	1.8		.05
SPIN	50.3	29.3	88.6	11.9	.01	.02
NU-6	42.9	14.6	61.4	13.8	.01	.05
Tracking	59.2	14.6	78.8	14.2	.01	.05

for F1 and duration, and somewhat higher with the CA processor for overall transmission and F2.

The remaining results presented in Fig. 2 and Table 3 are those for the segmental and open-set tests of the MAC battery, and for connected discourse tracking. The means of scores from the

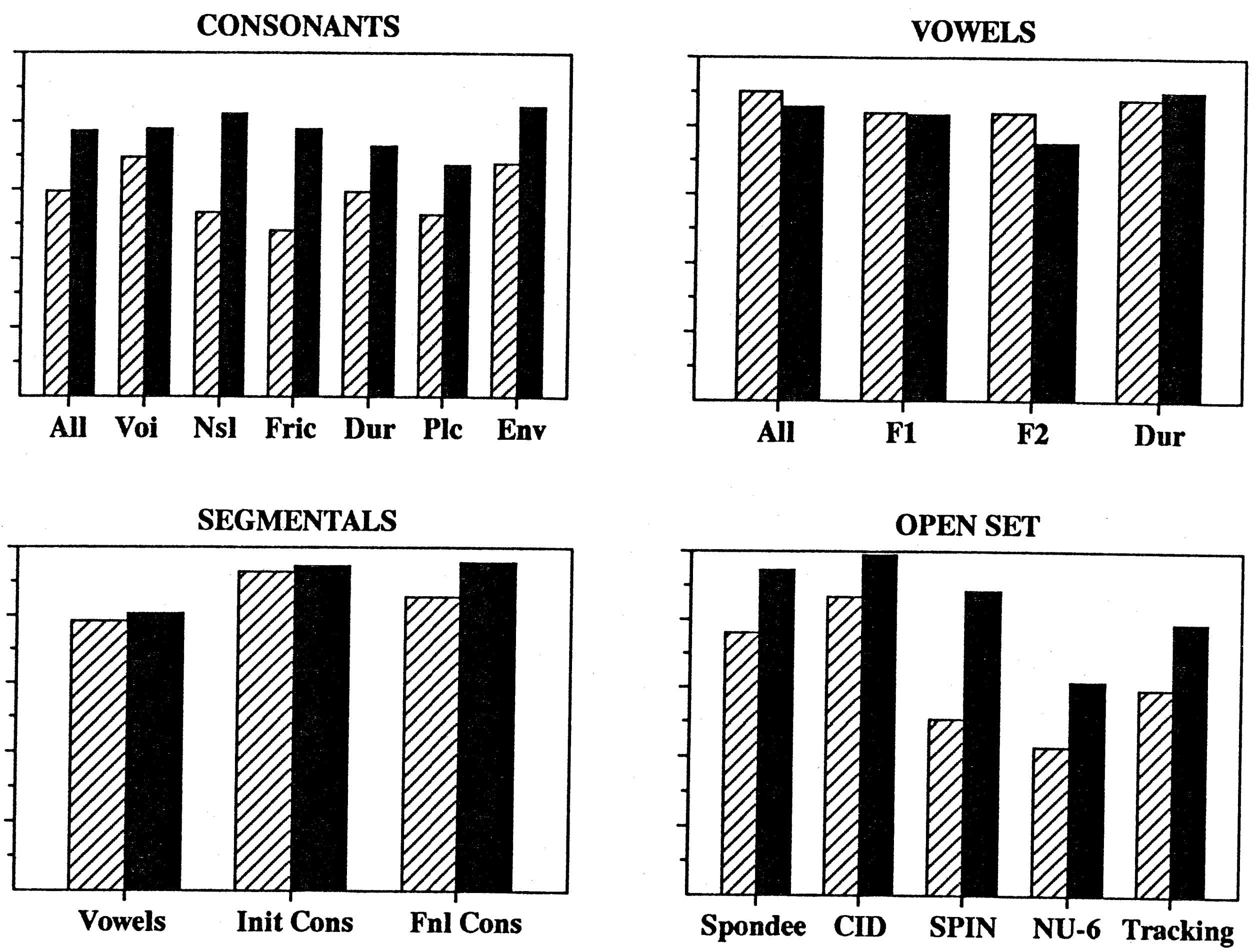

Fig. 2. Graphs of speech test results for the seven subjects. Scores for the CA processor are indicated by the striped bars, and those for the CIS processor by the solid bars. **(Top)** Relative information transfer of consonant and vowel features. The features include overall transmission (All), voicing (Voi), nasality (Nsl), frication (Fric), duration (Dur), place of articulation (Plc), envelope cues (Env), first formant frequency (F1), and second formant frequency (F2). Full scale corresponds to 100% information transfer. **(Bottom)** Average scores from the segmental and open-set tests. See text for abbreviations. Full scale corresponds to 100% correct for all tests except tracking, where full scale corresponds to 100 words per minute.

seven subjects are shown for all tests except tracking, where the means for six subjects are shown.

Results from the segmental tests all approximate perfect performance and mirror, to some extent, the results from the tests of vowel and consonant identification. In particular, the scores for the vowel test are indistinguishable, while the scores for the final consonant test demonstrate superiority of the CIS processor ($p < .02$). The scores for the initial consonant test do not favor either processor. However, ceiling effects may have masked a true difference between processors for that test (absolute scores are greater than 90% for both processors).

Finally, the open-set and tracking results demonstrate clear superiority of the CIS processor. Remarkable gains are found for all tests not subject to ceiling effects. The mean score for the SPIN test increases from 50.3 to 88.6% ($p < .01$); the mean score for the NU-6 test increases from 42.9 to 61.4% ($p < .01$); and the mean score for the tracking test increases from 59.2 to 78.8 words per minute ($p < .01$).

Individual Scores

An additional aspect of the open-set and tracking results is the pattern of increases for each subject. The individual scores for the CA and CIS processors are presented in Table 4. The scores for the CIS processor are those from the best tested variation of that processor. This variation was processor 1 for subjects SR2-6 and SR7, and processor 2 for subjects SR5 and SR8.

As indicated in the Table, every subject obtained a higher score, or repeated a score of 100% correct, for every test when the CIS processor was used instead of the CA processor. The increases across subjects are significant for spondee recognition ($p < .05$) and highly significant for recognition of the last word in the SPIN sentences ($p < .01$), recognition of the NU-6 words ($p < .002$), and the rate of speech tracking ($p < .02$). The increase for recognition of key words in the CID sentences is not significant, in part because the performance of several subjects is perfect or nearly so with both processors.

The overall pattern of scores in Table 4 was evaluated further with a two-way analysis of the variance (ANOVA), using the five tests and two processors as the factors. This analysis demonstrated highly significant differences among tests ($F[4,56] = 13.5$; $p < .0001$) and between processors ($F[1,56] = 34.1$; $p < .0001$), with no interaction between factors ($F[4,56] = 1.4$; $p > .2$).

Correlation Analyses

A final aspect of the results is shown in Table 5, a matrix of correlations among test scores for all subjects and both processors. The CIS processors are the ones listed first for each subject in Table 1.

As might be expected from the redundancy in assignments for consonant features (see Table A.3 and Wang and Bilger, 1973), high correlations are found among the transmission scores for those features. Also, overall transmission is highly correlated with all six consonant features ($r = .86$ or higher; $p < .001$), with especially high correlations observed for the features of nasality, frication, place, and envelope ($r = .94$ or higher).

Similarly, high correlations are found among transmission scores for the vowel features. In particular, a strong relationship is demonstrated between the scores for F1 and F2 ($r = .83$; $p < .001$). As with the consonant features, all vowel features are strongly correlated with overall transmission ($r = .80$ or higher; $p < .001$). Among these, F1 and F2 have higher correlations ($r = .88$ and $.93$ respectively) than duration ($r = .80$).

Table 4. Individual results from the open-set tests.

	Spondee		CID		SPIN		NU-6		Tracking	
Subject	CA	CIS	CA	CIS	CA	CIS	CA	CIS	CA	CIS
SR2	92	96	100	100	78	96	56	80	81	94
SR3	52	96	66	98	14	92	34	58	51	89
SR4	68	76	93	95	28	70	34	40	—	—
SR5	100	100	97	100	94	100	70	80	—	—
SR6	72	92	73	99	36	74	30	49	43	56
SR7	80	100	99	100	66	98	38	71	51	68
SR8	68	100	80	100	36	94	38	66	56	94

Correlations among scores for the segmental tests are either insignificant or barely significant (scores for the vowel and initial consonant test are weakly correlated: $r = .64$; $p < .02$). This suggests that these tests are relatively independent.

In contrast, correlations among the open-set tests generally are quite high. With the exception of the correlation for the NU-6 and CID tests ($r = .67$; $p < .01$), all correlations are .84 or higher ($p < .001$). This suggests that the open-set tests are not independent of each other and further that one or more of these tests might be omitted in future studies without any loss in information.

Examination of correlations among classes of tests demonstrates strong relationships between feature transmission scores from tests of consonant identification and scores from the tests of open-set recognition. With the exception of the CID sentence test, significant correlations are found for every consonant feature, with relatively high correlations for overall transmission ($r = .83$ or higher; $p < .001$), nasality ($r = .75$ or higher; $p < .002$), frication ($r = .71$ or higher; $p < .005$), and place ($r = .73$ or higher; $p < .005$). The lack of high correlations for the CID test may be a result of the very high scores obtained for that test across processors and subjects (Tables A.1 and A.2).

In addition to the high correlations between consonant features and open-set scores, high correlations are found between consonant features and the scores from the final consonant test of the MAC battery. Somewhat surprisingly, similar correlations are not found for the initial consonant test. As with the relative lack of correlations with the CID scores, though, this might be attributable to the uniformity of scores across subjects and processors for the initial consonant test.

As might be expected from the strong relationship between consonant features and scores from the final consonant test, the latter scores also are predictive of outcomes on the open-set tests. The correlations are .69 or higher ($p < .01$), with especially high correlations found for the Spondee

Table 5. Correlations among test results for all seven subjects and both processors. Correlation coefficients of .53, .61, .66, .70, .75 and .78 are significant at p < .05, .02, .01, .005, .002 and .001 respectively. Correlation coefficients that are significant at p < .001 are highlighted with **bold face** type.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Consonants																	
1. Overall																	
2. Voice	**.84**																
3. Nasal	**.96**	.73															
4. Fric	**.94**	**.84**	**.89**														
5. Dur	**.86**	.65	**.80**	.74													
6. Place	**.94**	**.78**	**.87**	**.85**	**.90**												
7. Envel	**.95**	**.89**	**.93**	**.92**	.74	**.83**											
Vowels																	
8. Overall	.42	.38	.55	.27	.38	.40	.50										
9. F1	.52	.50	.58	.44	.37	.46	.61	**.88**									
10. F2	.35	.33	.45	.18	.33	.27	.46	**.93**	**.83**								
11. Dur	.56	.50	.69	.46	.47	.51	.64	**.80**	.62	.62							
Segmentals																	
12. Vowel	.61	.64	.54	.40	.62	.66	.60	.54	.49	.46	.64						
13. Init Cons	.66	.59	.65	.48	.70	.55	.70	.55	.51	.62	.55	.64					
14. Fnl Cons	**.80**	.65	**.80**	**.81**	.52	.60	**.82**	.13	.20	.09	.47	.32	.55				
Open Set																	
15. Spondee	**.86**	.68	**.86**	**.84**	.69	.76	**.82**	.32	.42	.22	.53	.51	.63	**.85**			
16. CID	.70	.53	.70	.71	.52	.57	.69	.27	.37	.28	.37	.38	.52	.69	**.84**		
17. SPIN	**.83**	.60	**.84**	**.82**	.66	.73	.77	.27	.31	.15	.56	.44	.54	**.87**	**.97**	**.84**	
18. NU-6	**.79**	.60	.75	.71	**.79**	**.81**	.66	.24	.19	.12	.49	.59	.61	.69	**.87**	.67	**.89**

and SPIN tests (r = .85 and .87 respectively; p < .001).

Unlike the high correlations found for consonant features and scores from the open-set tests, only weak or insignificant correlations are observed for vowel features and those scores. Similarly, scores for the vowel test of the MAC battery are not generally predictive of the open-set scores. The only exception is a weak correlation between scores for the vowel and NU-6 tests (r

= .59; $p < .05$).

In summary, the correlation results appear to reflect the average scores presented in Fig. 2 and Table 3. Large differences in scores between processors are found for most consonant features, most open-set tests, and the final consonant test of the MAC battery. The correlation results show that these differences between processors covary across subjects. The remaining tests produce similar scores for the two processors (vowel features, vowel and initial consonant tests of the MAC battery, and, to some extent, the CID sentence test), and therefore those scores are not predictive of each other or of scores from the majority of open-set tests.

Discussion

The CA and CIS processing strategies were compared in tests with seven subjects implanted with the Symbion electrode array. Every subject obtained a higher score, or repeated a score of 100% correct, for all five open-set tests when the CIS processor was used instead of the CA processor. In addition, significant gains in the transmission of consonant information were demonstrated for the CIS processor. Performances on tests of vowel identification, and on the vowel and initial consonant tests of the MAC battery, were similar for the two processors. Finally, scores for the open-set tests were highly correlated (across subjects and processors) with transmission scores for consonant features.

The absolute levels of performance obtained with the CIS processor exceed by wide margins the previous levels reported in the open literature for any cochlear implant patient, using any type of device. The highest previously-reported score for the NU-6 test, for example, was 60% correct [Dorman et al., 1989]. Four of the seven subjects in present series exceeded this previous record, and two of the subjects had scores of 80% correct. This latter score is in the range of scores obtained by people with mild-to-moderate hearing losses when taking the same test [Bess and Townsend, 1977; Goetzinger, 1978]. Also, most scores for the remaining tests are near the respective upper scale limits: four subjects had scores of 96% or higher for the spondee test; all seven subjects had scores of 95% or higher for the CID test; five subjects had scores of 92% or higher for the SPIN test; and three of five tested subjects had tracking rates of 89 wpm or higher. Indeed, scores of 100% correct were not uncommon for the spondee and CID tests, and two subjects had tracking scores of 94 wpm. These scores are indistinguishable from those obtained by control subjects with normal hearing (e.g., four subjects with normal hearing took the same tracking test and got scores of 94, 94, 96 and 97 wpm; see Owens and Raggio, 1987).

The overall pattern of increases in open-set performance is even more compelling when one considers the large disparity in experience the subjects had with the two processors. Each subject had multiple years of daily experience with the CA processor at the time of our tests, but had only 15 minutes of experience (with informal conversation) before formal evaluation of the CIS processor. In previous studies using within-subject controls, such differences in experience have strongly favored the processor with the greatest duration of use [Dowell et al., 1987; Tyler et al., 1986].

Collectively the present findings show that close approximations to normal levels of speech recognition are possible with multichannel cochlear implants. In addition, they demonstrate clear superiority of the CIS processor for identification of consonants and for open-set recognition of words and sentences. The unprecedented performance obtained with the CIS processor offers new hope for implant recipients.

Acknowledgements

We are pleased to acknowledge the collaboration of Bob Wolford in the studies with all seven subjects, and of Don Eddington and Bill Rabinowitz in the studies with subject SR2. We also are indebted to Michael Dorman, Richard Tyler, Mary Lowder, and Korine Dankowski for their help in identifying the subjects for this series. Finally, we are most grateful for the time and interest contributed by each of the subjects.

References

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

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

Dorman MF, Hannley MT, Dankowski K, Smith L, McCandless G: Word recognition by 50 patients fitted with the Symbion multichannel cochlear implant. Ear Hear 1989, 10:44-49.

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

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

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

Goetzinger CP: Word discrimination testing. In Katz J (ed): *Handbook of Clinical Audiology, Second Edition.* Baltimore: Williams & Wilkins Company, 1978, pp 149-158.

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

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

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

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

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

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

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

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

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

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

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

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

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

Appendix

Table A.1. Scores from tests with the CA processor.

	Subject						
	SR2	SR3	SR4	SR5	SR6	SR7	SR8
Consonants							
1. Overall	67.7	50.1	63.3	80.1	68.9	77.2	67.0
2. Voice	77.9	53.7	67.1	83.0	75.9	75.8	61.1
3. Nasal	59.0	18.8	38.2	92.2	61.3	89.0	64.1
4. Fric	54.4	28.5	40.2	74.8	57.6	74.8	30.5
5. Dur	56.8	50.4	56.3	78.5	55.6	64.7	60.9
6. Place	55.0	45.4	48.8	67.2	51.7	67.2	54.4
7. Envel	73.8	43.2	64.9	86.2	78.1	84.5	71.3
Vowels							
8. Overall	96.7	77.4	87.0	97.4	95.2	100.0	97.4
9. F1	100.0	53.3	84.3	89.4	100.0	100.0	90.8
10. F2	89.7	50.8	90.1	96.4	93.4	100.0	96.4
11. Dur	100.0	74.4	69.3	100.0	90.8	100.0	100.0
Segmentals							
12. Vowel	88.3	70.0	75.0	76.7	76.7	76.7	85.0
13. Init Cons	93.8	82.8	93.8	100.0	92.2	89.1	96.9
14. Fnl Cons	86.5	75.0	80.8	96.2	88.5	88.5	84.6
Open Set							
15. Spondee	92.0	52.0	68.0	100.0	72.0	80.0	68.0
16. CID	100.0	66.0	93.0	97.0	73.0	99.0	80.0
17. SPIN	78.0	14.0	28.0	94.0	36.0	66.0	36.0
18. NU-6	56.0	34.0	34.0	70.0	30.0	38.0	38.0
19. Tracking	81.0	51.0	53.0	73.0	43.0	51.0	56.0

Table A.2. Scores from tests with the CIS processor.

	Subject						
	SR2	SR3	SR4	SR5	SR6	SR7	SR8
Consonants							
1. Overall	91.9	74.5	63.2	95.0	83.0	83.8	90.0
2. Voice	94.7	69.2	57.2	100.0	87.8	61.6	92.6
3. Nasal	100.0	68.8	54.7	100.0	85.5	96.1	100.0
4. Fric	87.9	63.9	56.0	100.0	88.9	74.8	97.2
5. Dur	92.4	55.1	46.7	96.7	81.4	92.2	68.2
6. Place	87.8	58.4	39.4	94.5	66.7	79.5	84.1
7. Envel	97.5	75.7	69.3	100.0	91.3	78.5	95.3
Vowels							
8. Overall	93.1	79.9	84.7	95.3	84.4	93.9	91.0
9. F1	82.2	63.6	75.3	100.0	82.2	94.7	100.0
10. F2	90.7	59.8	74.3	87.2	74.4	86.6	78.5
11. Dur	100.0	84.3	87.4	100.0	87.4	90.8	90.8
Segmentals							
12. Vowels	91.7	76.7	66.7	86.7	80.0	80.0	80.0
13. Init Cons	98.4	90.6	90.6	98.4	96.9	95.3	92.2
14. Fnl Cons	98.1	100.0	92.3	96.2	96.2	90.4	96.2
Open Set							
15. Spondee	96.0	96.0	76.0	100.0	92.0	100.0	100.0
16. CID	100.0	98.0	95.0	100.0	99.0	100.0	99.5
17. SPIN	96.0	92.0	70.0	100.0	74.0	98.0	90.0
18. NU-6	80.0	58.0	40.0	70.0	49.0	71.0	62.0
19. Tracking	94.0	89.0	—	81.0	56.0	68.0	85.0

Table A.3. Assignment of consonant features.

Consonant	Voicing	Nasality	Frication	Duration	Place	Envelope
m	2	2	1	1	1	4
n	2	2	1	1	2	4
f	1	1	2	1	1	3
v	2	1	2	1	1	2
s	1	1	2	2	2	3
z	2	1	2	2	2	2
ʃ	1	1	2	2	3	3
ʒ	2	1	2	1	1	2
p	1	1	1	1	1	1
b	2	1	1	1	1	2
t	1	1	1	1	2	1
d	2	1	1	1	2	2
g	2	1	1	1	4	2
k	1	1	1	1	4	1
dʒ	2	1	2	1	3	2
l	2	1	1	1	2	4

Table A.4. Assignment of vowel features.

Vowel	F1	F2	Duration
i	1	1	1
ɔ	2	2	1
ɛ	2	1	2
u	1	2	1
I	1	1	2
U	1	2	2
ʌ	2	3	2
æ	2	1	1

Table A.5. Consonant matrix for the CA processor.

	Response															
Stimulus	m	n	f	v	s	z	ʃ	ð	p	b	t	d	g	k	dʒ	l
m	122	45		4				6		5		4	1			20
n	3	189		1	1	1		3		1	1	6			1	2
f	1	119	13	4	1		29	14	11		4	7	2		1	
v	13	3	8	60	1	5		37	2	58		3	2			15
s	1	3	11	2	138	17	8	5	4	2	1	3	9	1		
z	2	7	4	23	5	123	1	18		1		4	8	4		8
ʃ					2		202	1								
ð	13	9	2	45	1	16		67		25		10	6			10
p	1		6	2	2	1		5	136	1	41			12		
b	9	3		17	1			7	3	156		3	2			6
t		1	2		1		1	1	15		138	1		46		
d		4		5		3		6		4	1	139	43	1	1	
g	1	14	3	9		2		8		2		21	142	3		3
k		2	2					2	33		40	2		128		
dʒ												2			203	1
l	16	43		3				4		1		4		1		137

Table A.6. Consonant matrix for the CIS processor.

	Response															
Stimulus	m	n	f	v	s	z	ʃ	ʒ	p	b	t	d	g	k	dʒ	l
m	116	25						1								3
n	10	130														4
f			116	13	3	2	1	10								
v	7	3	8	80				29		12		2				4
s			15	1	108	5	9	5	2							
z			4	9		114		16		1		2				
ʃ					11	1	133									
ʒ	2	1	3	41		2		88	1	1		1	1		3	
p								1	105		25			5	9	
b			1	3				5		131		5				
t					3		1				122			7	12	
d										4		113	25		2	
g												4	141			
k									7		5			131	1	
dʒ											2				143	
l	11	7										1				127

Table A.7. Vowel matrix for the CA processor.

	Response							
Stimulus	i	ɔ	ɛ	u	I	U	ʌ	æ
i	132							
ɔ		131						1
ɛ			125		7			
u				127		5		
I	2		1		128		1	
U			6	1	10	116		
ʌ		1	1			9	120	2
æ			1			1	3	127

Table A.8. Vowel matrix for the CIS processor.

	Response							
Stimulus	i	ɔ	ɛ	u	I	U	ʌ	æ
i	123			2	1			
ɔ		106					6	14
ɛ			116		4		6	
u	1			123	1	1		
I			9		116	1		
U			1	2		119	4	
ʌ			8	1	2	4	111	
æ		4	2					120

III. Plans for the Next Quarter

Our plans for the next quarter include the following:

1. Meet with principals at Richards Medical in Memphis, to discuss possible implementation of a portable CIS processor for use by patients using the Ineraid/Richards cochlear prosthesis (May 31).
2. Present project results at the *Second International Cochlear Implant Symposium*, to be held in Iowa City, IA, June 4-8, 1990.
3. Conduct studies with second Auditory Brainstem Implant (ABI) patient, in collaboration with investigators from the House Ear Institute (July 10-14).
4. Continue studies with Symbion patients (July 30 through August 10, and August 20-24).
5. Continue preparation of manuscripts for publication.

Appendix 1

Summary of Reporting Activity for the Period of

February 1 through April 30, 1990

NIH Contract N01-DC-9-2401

The following presentations were made in the last quarter of project work. The abstract for the presentation made at the ARO meeting is reproduced on the next two pages.

Finley, C.C. and B.S. Wilson: Spiral ganglion cell body effects on neural response latency in the electrically stimulated cochlea. Abstracts of the 13th Midwinter Research Conference, Association for Research in Otolaryngology, February 4-8, 1990, pp. 331-332.

Wilson, B.S.: Recent advances in the design of cochlear prostheses. Invited presentation given at Richards Medical, Memphis, TN, February 5, 1990.

Wilson, B.S.: Design of cochlear prostheses. Invited paper presented in the special session on "Cochlear Implants in Children," AAAS Meeting, New Orleans, February 15-20, 1990.

Wilson, B.S.: Comparison of Compressed Analog, Interleaved Pulses, and Supersampler Processors for Multichannel Cochlear Prostheses. Invited presentation given at the NIH site for the University of Iowa's Program Project Grant on cochlear prostheses, Iowa City, IA, April 17, 1990.

In addition to the above presentations, Wilson served as a member of the Science Advisory Council for the House Ear Institute (Los Angeles, April 18-19).

explantation, and reimplantation conditions.

Thirty subjects, in five experimental groups, were studied for four months, including: 1) implantation with a single wire ball tip electrode, 2) implantation with single wire ball tip electode followed by explantation at two months, 3) implantation with a single wire ball tip electrode followed by explantation at two months, and reimplantation with a new single wire ball tip electrode, 4) implantation with a single wire ball tip electrode with explantation at two months, and reimplantation with a silastic carrier simulating a multichannel device, and 5) implantation with a silastic carrier followed by explantation at two months. In the case of individual animals in each group, occasional organ of Corti damage was observed in the vicinity of the implant with some associated spiral ganglion cell degeneration. No significant differences in the average pathology across the experimental groups were observed. Thus, explantation, or explantation with reimplantation did not constitute an additional significant pathological risk to implantation alone.

(Work supported by NIH grant NS 21440 and FDA Contract 223-87-6028.)

376 SPIRAL GANGLION CELL BODY EFFECTS ON NEURAL RESPONSE LATENCY IN THE ELECTRICALLY STIMULATED COCHLEA.

*C. Finley, B. Wilson, Neuroscience Program Office, Research Triangle Inst., Research Triangle Park. N.C. 27709 and Dept. of Surgery, Div. of Otolaryngology, Duke Univ. Medical Ctr., Durham, NC

Single neuron studies of cat primary fiber responses to electrical stimulation have demonstrated an abrupt latency reduction in the neural response as stimulation intensity is increased (van den Honert and Stypulkowski, Hear Res. 29:207-222, 1987; Javel *et al.*, Ann Otol Rhinol & Laryngol, 96:(Suppl 128), 26-30, 1987). The magnitude of this latency shift is estimated to be 100 to 300 μsec. It has been proposed that the latency reduction is due to a shift in the location of the initial site of excitation along the fiber. Estimates of the distance between the proposed sites of excitation, assuming typical, uniform conduction velocities (5-20 m/sec) for myelinated fibers, are too large for the anatomical spacing of the possible nodal sites along cochlear neurons.

A possible explanation for this discrepancy is that the electrically stimulated fibers have a nonuniform conduction velocity between the two sites of excitation that produces a significant conduction delay. We propose that this conduction delay is introduced by the propagation of the action potential across the ganglion cell body and that, consequently, the two sites of excitation are separated by the presence of the cell body.

This hypothesis has been explored using an integrated field-neuron model of electrical stimulation (Finley and Wilson, ARO Abs. 8:105-106, 1985). Electrical field estimates are based on a three-dimensional, finite-element, field model of the cochlea (Finley *et al.*, Models of the Electrically Stimulated Cochlea, Miller and Spelman (eds.), Chap. 5, in press). Neural responses are

modeled using a modified McNeal model in which all node positions are active. The presence of the ganglion cell is modeled by approximating the cell body as a series of nine coaxially-aligned cylinders of varying length, diameter and wall thickness. The passive electrical characteristics of the model between two node sites are then adjusted based on the equivalent electrical analog of each cylinder.

This model has been used to evaluate the effects of the presence of the cell body on the responsiveness of a fiber to an extracellular electrical field. An important observation is that the cell body introduces a significant conduction delay in the central propagation of action potentials. This delay is sufficient to account for the latency shift phenomenon described above. This observation also is consistent with the timing characteristics of spike doublets recorded from the spiral ganglion during normal spontaneous activity. The specific location of excited nodes on either side of the cell body is not predicted by the present investigations.

(Supported by NIH Contract N01-NS-9-2401, "Speech Processors for Auditory Prostheses", Neural Prosthesis Program.)

377 ANALYTICAL MODEL FOR PASSIVE ELECTROTONUS AND ELECTRICAL STIMULATION OF MAMMALIAN MYELINATED FIBERS.

*J.T. Rubinstein, Dept. of Otolaryngology and Cochlear Implant Research Lab., Mass. Eye and Ear Infirmary, Boston, MA 02114

A passive cable model for myelinated nerve fibers has been developed using the theory of wave propagation in periodic structures. To the maximal extent possible, the model parameters are derived from published measurements of mammalian peripheral and auditory nerve. Nodal electrical properties have been obtained from the voltage-clamp literature on mammalian peripheral nerve. Internodal electrical properties are obtained from studies of frog fibers. Anatomic data are obtained from electron microscopic studies of feline auditory and peripheral nerve and HRP injections of single feline auditory nerve fibers.

The model permits simple calculation of length constants for a given fiber diameter and signal frequency. It predicts the linear relation between fiber diameter and conduction velocity that has been observed for myelinated fibers. It demonstrates the differences that should be observed between four different models of the node-paranode region.

When the model is stimulated by an extracellular electrode, it permits calculation of the passive membrane polarization and demonstrates several effects of source distance. If the stimulating electrode is electrically "far," the model behaves qualitatively similar to an unmyelinated fiber (Rubinstein and Spelman, Biophys. J. 54:975-981, 1988). If the source is electrically "near," the model predicts stop- and pass-bands in the spatial frequency domain whose amplitude is dependent on stimulation frequency. The characteristics of the stop-bands permit analysis of the assumption that myelin is a perfect insulator. This assumption is implicit in virtually all non-