X Original sent to printing

Fifth Quarterly Progress Report

May 1 through July 31, 1990

NIH project N01-DC-9-2401

"Speech Processors for Auditory Prostheses"

Prepared by

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

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

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. Completion 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 second Auditory Brainstem Implant (ABI) patient, to evaluate single-channel coding strategies for use with that implant.
3. Presentation of project results at the *Second International Cochlear Implant Symposium*, held in Iowa City, IA, June 4-8.
4. Initial development of a portable speech processor, based on the Motorola DSP 56001, for implementation of the CIS strategy in field tests.
5. Continued preparation of manuscripts for publication.

In this report we present results from the final subject in the Symbion series (point 1 above). Compared to the 7 subjects described in our last progress report, this subject had relatively poor performance with his clinical, compressed analog (CA) processor. Thus, studies with him provided the opportunity to compare the CA and CIS strategies with a more typical patient than the star performers of *QPR 4*. As described in detail in section II of this report, use of the CIS strategy produced large improvements in speech recognition scores for this last subject. Recognition of spondee words increased from 40% correct with the CA strategy to 60% correct with the CIS strategy; recognition of key words in the CID sentences increased from 25% to 70%; recognition of the final words in the high-context SPIN sentences increased from 2% to 30%; and recognition of the NU-6 monosyllabic words increased from 6% to 32%. In relative terms, these gains are greater than those observed for any of the 7 subjects with high levels of initial performance with the CA strategy.

Complete descriptions of our work with ABI patients (point 2), and of the development of the 56001-based portable processor (point 4), will be presented in future reports.

II. Further Evaluation of the *Continuous Interleaved Sampler* Strategy

In previous studies, reviewed in *QPR 4*, we compared the compressed analog (CA) and continuous interleaved sampler (CIS) strategies in tests with seven subjects who had excellent performance with their CA processors. As described in *QPR 4*, each of those subjects obtained a higher score, or repeated a score of 100% correct, for all open-set tests when the CIS processor was used instead of the CA processor. In addition, significant gains in the transmission of consonant features were demonstrated with the CIS processor. Performances on tests of vowel identification, and on the vowel and initial consonant tests of the Minimal Auditory Capabilities (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.

In the present study, we compared the CA and CIS strategies in tests with a subject who had relatively poor performance with his CA processor. The purpose of these latter tests was to determine if such a subject (SR1) might also benefit from application of the CIS strategy.

Methods

Subject SR1 lost his hearing at age 8 as a consequence of spinal meningitis. He was implanted at age 40, and first studied by us at age 43 [*QPR 1*; Wilson et al., in press]. The first studies were done in the spring of 1989, and the present studies in the summer of 1990. Both the long duration of deafness and the etiology of meningitis have been identified as prognosticators of poor performance with current cochlear prostheses [e.g., Dorman et al., 1990].

With the exception of connected discourse tracking with hearing alone, all tests described in *QPR 4* for the initial seven subjects (SR2-8) also were conducted with SR1. As with the previous subjects, the CIS processor used for SR1 had a staggered order of channel updates, a relatively high rate of stimulation on each channel (833 Hz), and a relatively high corner frequency for the RMS energy detectors (400 Hz). Parameters for the CIS processors used by all eight subjects are presented in Table 1.

We note that the selection of the "best" CIS processor was somewhat problematic for SR1 because his scores on the consonant identification test were more variable from block to block than those for any of the initial seven subjects. Whether this is a general characteristic of subjects with relatively low levels of performance with their implants, or a particular characteristic of SR1, remains to be seen in future studies.

Results

Results from Seven Subjects with High Levels of Initial Performance

Results obtained from tests with the first processors listed in Table 1 for subjects SR2-8 are shown Fig. 1. The top two panels show the information transmission (IT) scores for consonant and vowel features, and the bottom panels show average scores for the segmental and open-set tests of

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 $\mathit{pulse\ amplitude} = A\ x\ \mathit{log(RMS)} + k$, and the power-law transformation is of the form $\mathit{pulse\ amplitude} = A\ x\ \mathit{(RMS)}^p + k$, 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. Parameters for the subject of the present study, SR1, are highlighted with **boldface** type.

Subject	μs/ph	RMS rect /	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 ($p = 0.2$)
SR1	**34**	**HW**	**400**	**1200**	**6-3-1-5-2**[d]	**833**	**logarithmic**

[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 was fitted with a five-channel processor because stimulation of his sixth electrode produced transient sensations of head movements.

[d] **Subject SR1 was fitted with a five-channel processor, omitting electrode 4 because of its higher thresholds.**

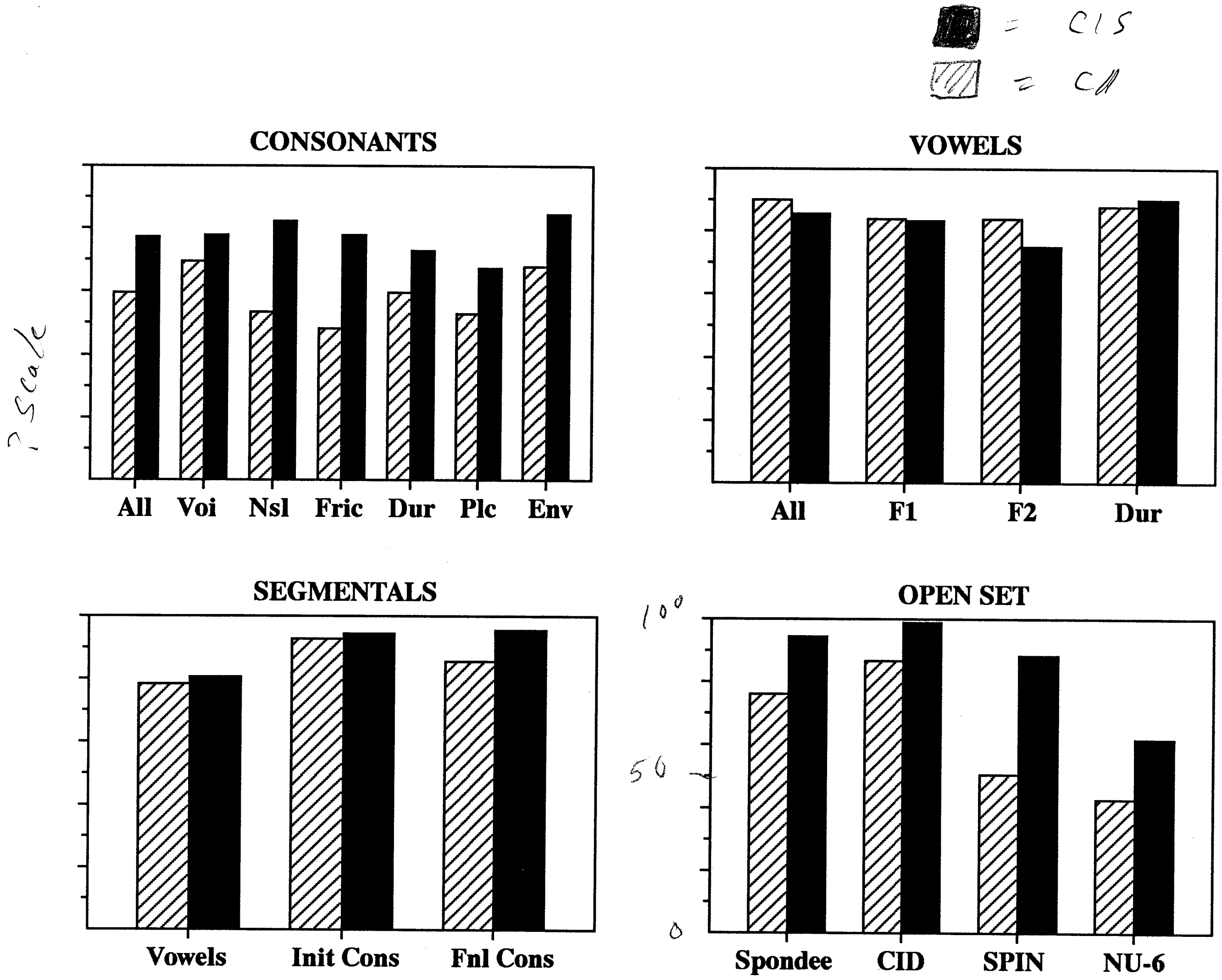

Fig. 1. Results from tests with the first seven subjects. Scores for the CA processor are indicated by the striped bars, and those for the CIS processor by the solid bars. **(Top)** Relative information transfer of consonant and vowel features. The features include overall transmission (All), voicing (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. Scales are from 0 to 100% correct.

the MAC battery.

The IT scores demonstrate large gains in the consonant features of overall transmission, nasality, frication, and place of articulation when the CIS processor is used instead of the CA processor. In addition, substantial increases are found for consonant duration and envelope cues. Finally, note that the absolute scores for all features except place exceed 70% when the CIS processor is used, and the scores for nasality and envelope each exceed 80%. The greatest strengths of the CIS processor are in the transmission of nasality, frication 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 for the transmission of vowel features are quite high for both processors and all features. Transmission scores are nearly identical for F1 and duration, and somewhat higher with the CA processor for overall transmission and F2.

Results from the segmental tests are also quite high for both processors. The scores for the final consonant test are significantly better for the CIS processor ($p < .02$). Scores for the vowel and initial consonant tests do not favor either processor, but in the latter case the test's sensitivity to differences was limited by scores exceeding 90% for both processors.

Finally, the open-set results demonstrate clearly better performance with the CIS processor. Remarkable gains are found for all tests not subject to ceiling effects. The increases across subjects are significant for spondee recognition ($p < .05$), recognition of the last word in the SPIN sentences ($p < .01$), and recognition of the NU-6 words ($p < .01$). The increase for recognition of key words in the CID sentences is not statistically significant. The performance of several subjects on that test is perfect or nearly so with both processors.

Results from One Subject with Low Levels of Initial Performance

Results from tests with subject SR1 are shown in Fig. 2. As mentioned before, this subject had relatively low levels of performance with his clinical CA processor.

The pattern of results in Fig. 2, for SR1, is remarkably similar to the pattern of results in Fig. 1, for the subjects with high levels of initial performance with their clinical devices. In particular, large gains are found in the transmission of consonant features and in the recognition of open-set material when the CIS processor is used instead of the CA processor. The increases in IT scores for subject SR1 (Fig. 2) mirror those for subjects SR2-8 (Fig. 1). Especially large increases in overall transmission and the transmission of nasality, frication and envelope information are seen in both sets of data. However, the magnitudes of the increases are larger for subject SR1.

As with the first seven subjects, SR1's scores for the transmission of vowel features are similar for the two processors. The previous pattern of slightly higher scores for overall transmission, F1 and F2 with the CA processor, and of a slightly higher score for the transmission of duration with the CIS processor, is repeated in the results obtained with subject SR1.

In addition, the statistically insignificant increases seen for the segmental tests with the first seven subjects become quite large for SR1 (Fig. 2).

Finally, the open-set results for SR1 again demonstrate clear improvement with the CIS processor. A large gain in the recognition of spondee words is realized with the CIS processor, and scores for the CID, SPIN and NU-6 tests are more than doubled with the use of that processor.

Individual Scores from All Eight Subjects

An additional aspect of the open-set and tracking results is illustrated in Table 2. Here, the

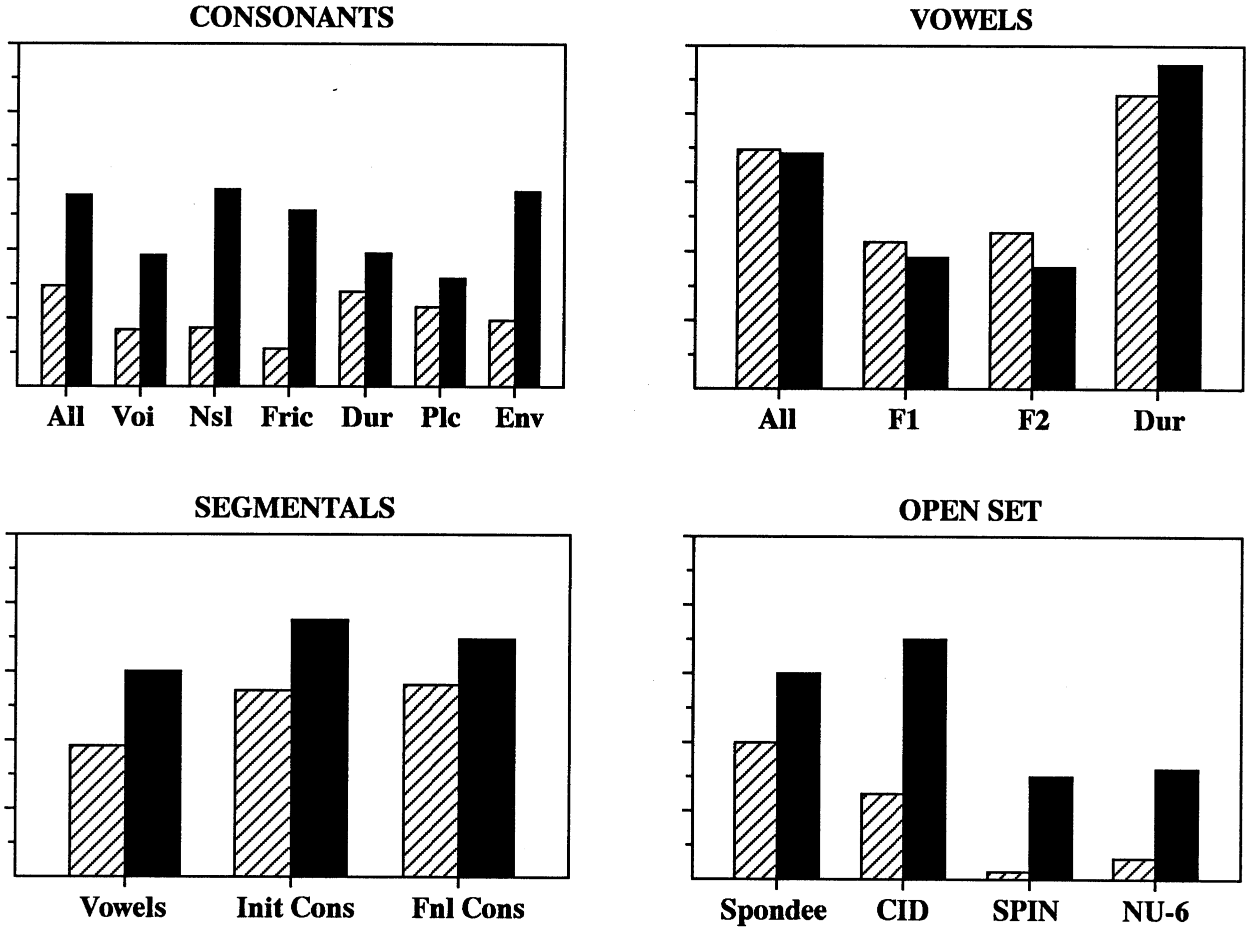

Fig. 2. Results from tests with subject SR1. Scores for the CA processor of the Symbion device are indicated by the striped bars, and scores for the CIS processor by the solid bars. Full scale in the upper panels corresponds to 100% information transfer, and full scale in the lower panels corresponds to 100% correct.

individual scores for the CA and CIS processors are presented for all eight subjects. The scores for the CIS processor are those from the best tested variation of that processor. This corresponds to the second processor listed for subjects SR5 and SR8, for whom two variations were included in Table 1.

As indicated in Table 2, 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 the recognition of key words in the CID sentences ($p < .05$), and highly significant for spondee recognition ($p < .01$), recognition of the last word in the SPIN sentences ($p < .01$), recognition of the NU-6 words ($p < .001$), and the rate of speech tracking

Table 2. Individual results from the open-set tests. Results for SR1 are highlighted with **boldface** type.

	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
SR1	**40**	**60**	**25**	**70**	**2**	**30**	**6**	**32**	—	—

($p < .02$).

The overall pattern of scores in Table 2 was evaluated further with a two-way analysis of the variance, using the five tests and two processors as the factors. This analysis demonstrated highly significant differences among tests ($F[4,64] = 8.51$; $p < .0001$) and between processors ($F[1,64] = 23.77$; $p < .0001$), with no significant interaction between factors ($F[4,64] = 0.66$; $p > .6$).

Discussion

The results obtained in the tests with subject SR1 show that a patient starting with low levels of performance with one processing strategy may receive large benefits from substitution of another strategy. In particular, use of the CIS rather than the clinical CA strategy moved subject SR1 from barely measurable open-set performance to high levels of open-set performance. Indeed, with the CIS strategy subject SR1 has scores that fall within the range of those obtained by the first seven subjects (SR2-8) with their CA processors. As noted in *QPR 4*, such scores are among the best previously recorded for cochlear implant users.

The present findings are especially encouraging in that most patients, unfortunately, have low levels of performance with their clinical devices. Substantial increases in those levels can bring such patients into the domain of useful open-set recognition with hearing alone, and may broaden the application of cochlear implants as a treatment for sensori-neural deafness. We plan to study additional subjects with low levels of initial performance, to assess the generality our findings with subject SR1.

Acknowledgements

We are pleased to acknowledge the collaboration of Bob Wolford, Don Eddington and Bill Rabinowitz in the studies with SR1. We also are indebted to the subject for his generous contribution of time and interest.

References

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.

Wilson BS, Lawson DT, Finley CC, Wolford RD: Coding strategies for multichannel cochlear prostheses. Am J Otol, in press.

III. Plans for the Next Quarter

Our plans for the next quarter include the following:

1. Presentation of project results at the *21st Neural Prosthesis Workshop* (October 17-19).
2. Continued preparation of manuscripts for publication.
3. Continued studies with Symbion patient MP (SR2), to evaluate parametric variations of the continuous interleaved sampler (CIS) and other processors.
4. Electric field mapping and electrophysiological studies with a patient with direct percutaneous access to her implanted UCSF/Storz electrode array.
5. Continued development of a portable processor based on the DSP 56001 device.

Appendix 1

Summary of Reporting Activity for the Period of

May 1 through July 31, 1990

NIH Contract N01-DC-9-2401

The following presentations were made in the last quarter of project work. An abstract for the first presentation is presented on the next page.

Wilson, B.S., C.C. Finley, D.T. Lawson and R.D. Wolford: A new processing strategy for multichannel cochlear implants. Presented at the *Second International Cochlear Implant Symposium*, Iowa City, IA, June 4-8, 1990.

Wilson, B.S.: Moderator: Session on Speech Processing. *Second International Cochlear Implant Symposium*, Iowa City, IA, June 4-8, 1990.

Shannon, R.V. (moderator), B.S. Wilson, D.K. Eddington, J. Walliker, B.E. Pfingst, J.F. Patrick and S. Rosen (panelists): Round table discussion on "Future Directions in speech Processing." *Second International Cochlear Implant Symposium*, Iowa City, IA, June 4-8, 1990.

A NEW PROCESSING STRATEGY FOR MULTICHANNEL COCHLEAR IMPLANTS

B Wilson, C Finley, D Lawson

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

Three strategies for representing speech information with a multichannel cochlear implant were evaluated in tests with a deaf subject. The results demonstrated large differences among strategies and unprecedented levels of performance with one of the strategies. This best strategy presented brief pulses in immediate succession across six channels, with the pulse amplitudes for each channel reflecting the energy in a corresponding frequency band. The high rate of stimulation on each channel was designed to improve the representation of temporal events, while the use of nonsimultaneous pulses was designed to increase the salience of channel cues through elimination of current summation between channels.

Studies to extend comparisons of these processing strategies to a population of implant patients are in progress; results from additional subjects will be presented at the meeting.