Department of
Veterans Affairs

Journal of Rehabilitation Research
and Development Vol. 30 No. 1, 1993
Pages 110–116

Design and evaluation of a continuous interleaved sampling (CIS) processing strategy for multichannel cochlear implants

Blake S. Wilson, PhD; Charles C. Finley, PhD; Dewey T. Lawson, PhD; Robert D. Wolford, MS; Mariangeli Zerbi, MSEE
Neuroscience Program, Research Triangle Institute, Research Triangle Park, NC 27709;
Division of Otolaryngology, Duke University Medical Center, Durham, NC 27710

Abstract—Two approaches for representing speech information with multichannel cochlear prostheses are being compared in tests with implant patients. Included in these studies are the *compressed analog* (CA) approach of a standard clinical device and research processors utilizing *continuous interleaved sampling* (CIS). Initial studies have been completed with nine subjects, seven of whom were selected on the basis of excellent performance with the Ineraid clinical processor, and the remaining two for their relatively poor performance with the same device. The tests include open-set recognition of words and sentences. Every subject has obtained a higher score—or repeated a score of 100% correct—on every test when using a CIS processor. These results are discussed in terms of their implications for processor design.

Key words: *cochlear prosthesis, deafness, hearing, speech perception, speech processing.*

INTRODUCTION

Recent studies in our laboratory have focused on comparisons of *compressed analog* (CA) and *continuous interleaved sampling* (CIS) processors (1,2,3). Both use multiple channels of intracochlear electrical stimulation, and both represent waveforms or envelopes of speech input signals. No specific features of the input, such as the fundamental or formant frequencies, are extracted or explicitly represented. CA processors use continuous analog signals as stimuli, whereas CIS processors use pulses. The CA approach is used in the widely applied Ineraid device (4,5) and in the now-discontinued UCSF/Storz device, with some differences in details of processor implementation (6). Wearable devices capable of supporting the CIS approach are just becoming available for use in clinical settings.

To date, we have completed initial studies of nine subjects—seven of whom were selected for their high levels of speech recognition with the Ineraid CA processor, and two who were selected for their relatively poor performances with that processor. The "high performance" subjects were representative of the best results when any commercially available implant system is used (2). Equivalent studies have been begun but not yet completed with two additional patients in the "poor performance" group (7).

This paper will briefly review the previously published results for the seven subjects in the high performance group and present preliminary results for the first two subjects from the poor performance group.

PROCESSING STRATEGIES

The designs of CA and CIS processors are illustrated in **Figure 1** and **Figure 2**. In CA proces-

---

Address all correspondence and requests for reprints to: Blake S. Wilson, Neuroscience Program, Research Triangle Institute, Research Triangle Park, NC 27709.

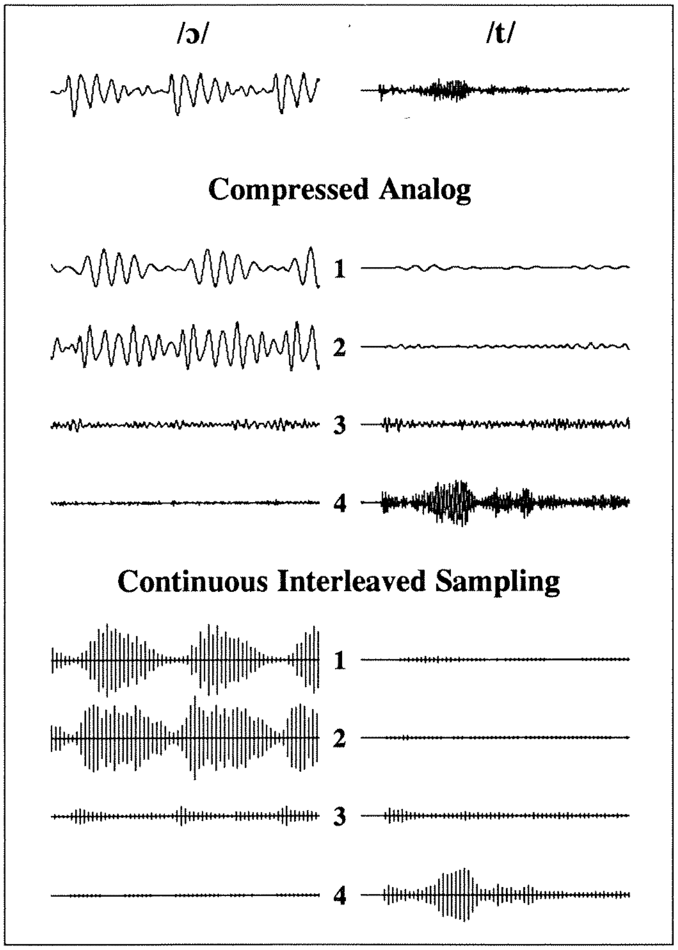

Figure 1.
Waveforms produced by simplified implementations of CA and CIS strategies. The top panel shows preemphasized (6dB/octave attenuation below 1.2 kHz) speech inputs. Inputs corresponding to a voiced speech sound ("aw") and an unvoiced speech sound ("t") are shown in the left and right columns, respectively. The duration of each trace is 25.4 ms. The remaining panels show stimulus waveforms for CA and CIS processors. The waveforms are numbered by channel, with channel 1 delivering its output to the apical-most electrode. To facilitate comparisons between strategies, only four channels of CIS stimulation are illustrated here. In general, five or six channels have been used for that strategy. The pulse amplitudes reflect the envelope of the bandpass output for each channel. In actual implementations the range of pulse amplitudes is compressed using a logarithmic or power-law transformation of the envelope signal. (From Wilson BS, et al. (2), with permission.)

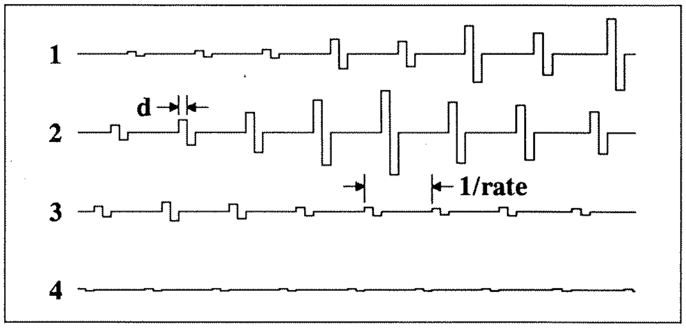

Figure 2.
Expanded display of CIS waveforms. Pulse duration per phase ("d") and the period between pulses on each channel ("1/rate") are indicated. The sequence of stimulated channels is 4-3-2-1. The total duration of each trace is 3.3 ms. (From Wilson BS, et al. (2), with permission.)

sors, a microphone signal varying over a wide dynamic range is compressed or restricted to the narrow dynamic range of electrically evoked hearing (8,9) using an automatic gain control. The resulting signal is then filtered into four contiguous frequency bands for presentation to each of four electrodes. As shown in **Figure 1**, information about speech sounds is contained in the relative stimulus amplitudes among the four electrode channels and in the temporal details of the waveforms for each channel.

A concern associated with this method of presenting information is that substantial parts of it may not be perceived by implant patients (10). For example, most patients cannot perceive frequency changes in stimulus waveforms above about 300 Hz (11). Thus, many of the temporal details present in CA stimuli probably are not accessible to the typical user.

In addition, the simultaneous presentation of stimuli may produce significant interactions among channels through vector summation of the electric fields from each electrode (12). The resulting degradation of channel independence would be expected to reduce the salience of channel-related cues. That is, the neural response to stimuli from one electrode may be significantly distorted, or even counteracted, by coincident stimuli from other electrodes.

The CIS approach addresses the problem of such channel interactions through the use of interleaved nonsimultaneous stimuli (**Figure 2**). Trains of balanced biphasic pulses are delivered to each electrode with temporal offsets that eliminate any overlap across channels. The amplitudes of the pulses are derived from the envelopes of bandpass filter outputs. In contrast with the four-channel clinical CA processors, five or six bandpass filters (and channels of stimulation) have generally been used in CIS systems to take advantage of additional implanted electrodes and reduced interactions among channels. The envelopes of the bandpass outputs are formed by rectification and lowpass filtering. Finally, the amplitude of each stimulus pulse is determined by a logarithmic or power-law transformation of the corresponding channel's envelope signal at that time. This transformation compresses each signal into the dynamic range appropriate for its channel.

A key feature of the CIS approach is its relatively high rate of stimulation on each channel. Other pulsatile strategies present sequences of interleaved pulses across electrodes at a rate equal to the estimated fundamental frequency during voiced speech and at a jittered or fixed (often higher) rate during unvoiced speech (13,14,15). Rates of stimulation on any one channel have rarely exceeded 300 pulses per second (pps). In contrast, the CIS strategy generally uses brief pulses and minimal delays, so that rapid variations in speech can be tracked by pulse amplitude variations. The rate of stimulation on each channel usually exceeds 800 pps and is constant during both voiced and unvoiced intervals. A constant high rate allows relatively high cutoff frequencies for the lowpass filters in the envelope detectors. With a stimulus rate of 800 pps, for instance, lowpass cutoffs can approach (but not exceed) 400 Hz without introducing aliasing errors in the sampling of the envelope signals at the time of each pulse. See Rabiner and Shafer for a complete discussion of aliasing and its consequences (16).

METHODS

Each of the nine subjects has been studied for a 1-week period during which: (*a*) basic psychophysical measures were obtained on thresholds and dynamic ranges for pulsatile stimuli; (*b*) various CIS processors (with different choices of processor parameters) were evaluated with preliminary tests of consonant identification; and, (*c*) performance with the best of the CIS processors and the clinical CA processor was documented with a broad spectrum of speech tests. Experience with the clinical processor exceeded one year of daily use for all subjects. In contrast, experience with the CIS processors was limited to no more than several hours before formal testing.

Tests

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

Processor Parameters

Each subject's own clinical device is used for the tests with the CA processor. As mentioned above, selection of parameters for the CIS processor is guided by preliminary tests of consonant identification. The standard four channels of stimulation are used for the clinical CA processors (4,5), whereas five or six channels have been used for the CIS processors. Additional parameters of the CIS processors are presented in **Table 1**. As indicated there, all CIS processors for the high performance subjects, SR2 to SR8, have had pulse durations of 102 μs/phase or less, zero delay between the sequential pulses on different channels, pulse rates of 817 pps or higher on each channel, and a cutoff frequency for the lowpass filters of 400 Hz or higher. The best processor for subject SR1 also fit this description, except that a delay of 172 μs was interposed between sequential pulses. The best processor for subject SR10 used long-duration pulses (167 μs/phase), paired with a relatively low rate of stimulation on each channel (500 pps) and a relatively low cutoff frequency for the lowpass filters (200 Hz).

Evaluation of Practice and Learning Effects

Because the tests with the CA processor preceded those with the selected CIS processor for each subject, we were concerned that practice or learning effects might favor the latter in comparisons of the two strategies. To evaluate this possibility, the CID and NU-6 tests were repeated with the CIS processor for five of the high performance subjects (subjects SR3, SR4, SR6, SR7, and SR8). A different recorded speaker and new lists of words and sentences were used. Practice or learning effects would be demonstrated by significant differences in the

Table 1.
Parameters of CIS processors.

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

Parameters include number of channels, pulse duration, the rate of stimulation on each channel (Rate), and the cutoff frequency of the lowpass integrating filters for envelope detection (Integrating Filter Cutoff). The subjects are listed in the chronological order of their participation in the present studies. SR2 through SR8 are the "high performance" subjects while SR1 and SR10 belong to the "low performance" group. Additional processor parameters may be found in References 3 and 7.

test/retest scores. Nevertheless, 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

The results from 1-week studies of each of the nine subjects are presented in **Table 2** and **Figure 3**. Scores for the high performance subjects are indicated by the light lines near the top of each panel in **Figure 3**, and scores for the two low performance subjects are indicated by the dark lines closer to the bottom of each panel. We note that low performance subject SR1 had participated in an earlier study not involving CIS processors (15). Results from his first week of testing with CIS processors are presented here. This is also true of high performance subject SR2, who has returned to the laboratory for many additional studies with various implementations of CIS processors (1). In those subsequent tests, SR2 has achieved even higher scores using a variety of six-channel CIS processors, with NU-6 percentages ranging from the high 80s to the low 90s.

As is evident from the figure, all nine subjects have scored higher, or repeated a score of 100 percent correct, on every test, when using a CIS processor. The average scores across subjects increased from 64 to 86 percent correct on the spondee test ($p < 0.01$), from 70 to 91 percent correct on the CID test ($p < 0.02$), from 39 to 76 percent correct on the SPIN test ($p < 0.001$), and from 34 to 54 percent correct on the NU-6 test ($p < 0.0002$).

Perhaps the most encouraging of these results are the improvements for the two low performance subjects. CA scores were low for SR1 and quite poor for SR10. Substitution of a CIS processor produced large gains in speech recognition for both subjects. Indeed, with the CIS processor SR1 has scores that fall within the ranges of CA processor scores that qualified subjects SR2 to SR8 as among the best performers with any clinical device.

Similarly, SR10 achieved relatively high scores with the CIS processor. The score on the spondee test increased from 0 to 56 percent correct, on the CID test from 1 to 55 percent correct, on the SPIN test from 0 to 26 percent correct, and on the NU-6 test from 0 to 14 percent correct. These increases were obtained with no more than several hours of aggregated experience with CIS processors, compared with more than a year of daily experience with the clinical CA processor.

Note that although these gains for SR10 are large, they are not atypical of results for the other subjects. His improvements follow the pattern of the

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

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

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

DISCUSSION

The findings presented in this paper demonstrate that the use of CIS processors can produce large and immediate gains in speech recognition for a wide range of implant patients. Indeed, the sensitivity of some of the administered tests has been limited by ceiling (or saturation) effects: five of the seven high performance subjects scored 96 percent or higher for the spondee test using CIS processors; all seven scored 95 percent or higher for the CID test; and five scored 92 percent or higher for the SPIN test. Scores for the NU-6 test, although not approaching the ceiling, were still quite high. The 80 percent score achieved by two of the subjects corresponds to the middle of the range of scores obtained by people with mild-to-moderate hearing losses when taking the same test (17,18).

The improvements are even more striking when one considers the large disparity in experience with the two processors. At the time of our tests each subject had 1 to 5 years of daily experience with the CA processor but only several hours over a few days with CIS. In previous studies involving within-subjects comparisons, such differences in experience have strongly favored the processor with the greatest duration of use (19,20,21).

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

An interesting aspect of the ongoing studies with low performance subjects, represented here by SR1 and SR10, is that the best CIS processors seem to involve parameters distinct from those of the best processors for subjects in the high performance group. The best processor for SR1 used short-duration pulses (34 μs/phase) presented at a relatively low rate (833 pps), and the best processor for SR10 used long-duration pulses (167 μs/phase) presented at an even lower rate (500 pps). The subjects in the high performance group, however, often obtained their best scores with processors tending to minimize pulse widths and maximize pulse rates (e.g., 31 μs/phase pulses presented at 2688 pps).

The use of such shorter pulses and higher rates allows representation of higher frequencies in the modulation waveform for each channel (i.e., the

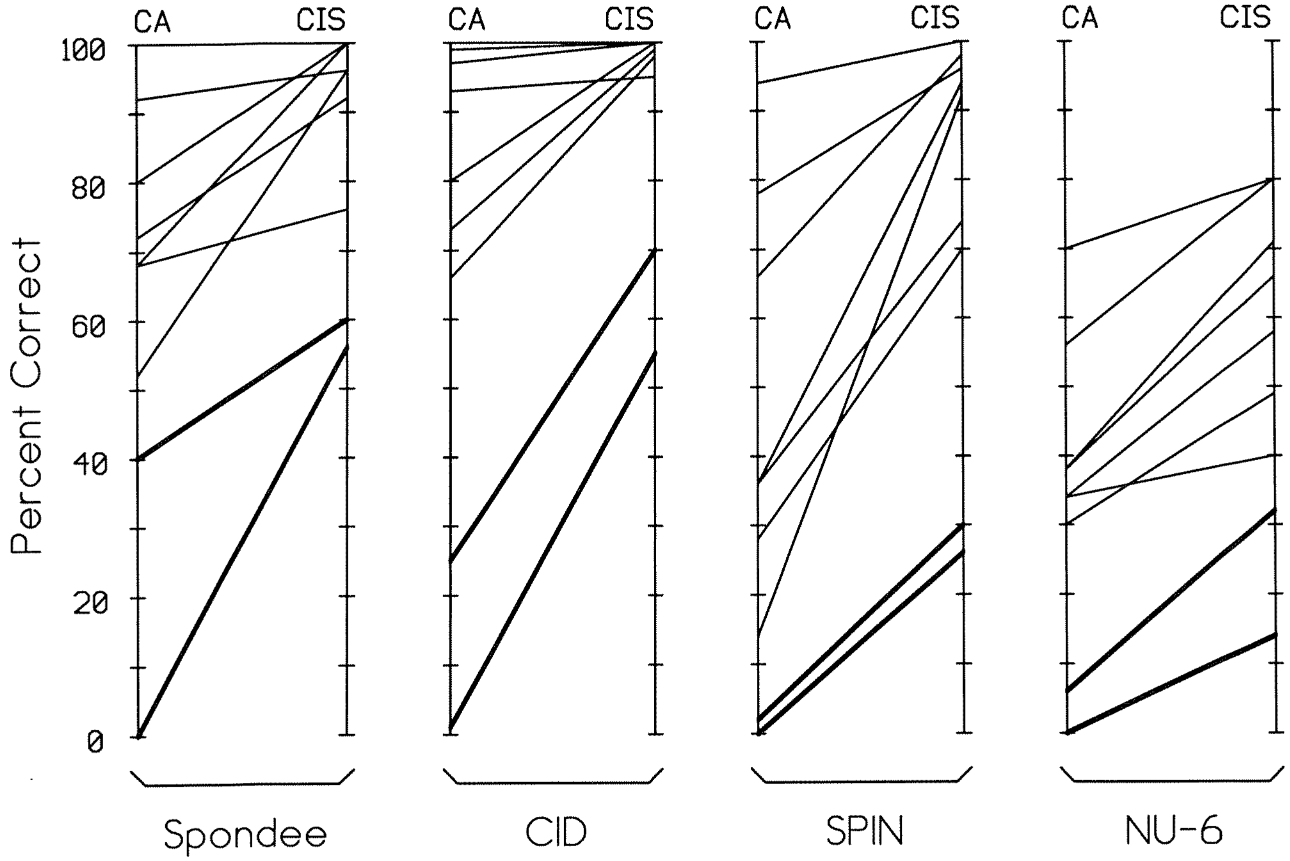

Figure 3.
Speech recognition scores for CA and CIS processors. A line connects the CA and CIS scores for each subject. Light lines correspond to the seven subjects selected for their excellent performance with the clinical CA processor, whereas the heavier lines correspond to the two subjects selected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

cut-off frequency of the lowpass filter in the envelope detectors for each channel may be raised to one-half of the pulse rate without introducing aliasing effects). In addition, the dynamic range (DR) of electrical stimulation—from threshold to most comfortable loudness—is a strong function of pulse rate and a weaker function of pulse duration (11,22). Large increases in DR are generally found with increases in pulse rates from about 400 pps to 2500 pps. Smaller increases often (but not always) are observed with increases in pulse duration (at a fixed rate of stimulation) from roughly 50 μs/phase to higher values (e.g., out to 200 μs/phase for practical CIS designs).

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

Another possible advantage of relatively low rates of stimulation is further reduction of channel interactions. Providing time between pulses on sequential channels can reduce the "temporal integration" component of channel interactions—a component produced by the accumulation of charge at neural membranes from sequential stimuli (12). Thus, use of time delays between short-duration pulses in the stimulation sequence across electrodes may reduce interactions. Alternatively, use of long-duration pulses with no time delay also might reduce temporal interactions in that a relatively long period still is realized between the excitatory phases of successive pulses.

Collectively, the present results indicate that: (*a*) the performance of at least some patients with poor clinical outcomes can be improved substantially with use of a CIS processor; (*b*) use of long-duration pulses produced large gains in speech test scores for one such subject; (*c*) use of short-duration pulses presented at a relatively low rate produced similar improvements in another such subject; and, (*d*) the optimal tradeoffs among pulse duration, pulse rate, interval between sequential pulses, and cutoff frequency of the lowpass filters seem to vary from patient to patient. Studies are underway to evaluate CIS processors for additional subjects in the low performance group and to investigate in detail the tradeoffs among processor parameters for subjects in both the low and high performance groups.

ACKNOWLEDGMENTS

We thank the subjects of the described studies for their enthusiastic participation. We also are pleased to acknowledge the important scientific contributions of Michael F. Dorman, Donald K. Eddington, William M. Rabinowitz, and Robert V. Shannon. This work was supported by NIH project N01-DC-9-2401, through the Neural Prosthesis Program.

REFERENCES

1. Lawson DT, Wilson BS, Finley CC. New processing strategies for multichannel cochlear prostheses. Prog Brain Res. In press.
2. Wilson BS, Finley CC, Lawson DT, Wolford RD, Eddington DK, Rabinowitz WM. Better speech recognition with cochlear implants. Nature 1991;352:236-8.
3. Wilson BS, Lawson DT, Finley CC. Speech processors for auditory prostheses. Bethesda (MD): National Institutes of Health, 1990. 4th Quarterly Progress Report, NIH project N01-DC-9-2401.
4. Eddington DK. Speech discrimination in deaf subjects with cochlear implants. J Acoust Soc Am 1980;68:885-91.
5. Eddington DK. Speech recognition in deaf subjects with multichannel intracochlear electrodes. Ann NY Acad Sci 1983;405:241-58.
6. Merzenich MM, Rebscher SJ, Loeb GE, Byers CL, Schindler RA. The UCSF cochlear implant project. State of development. Adv Audiol 1984;2:119-44.
7. Wilson BS, Lawson DT, Finley CC, Zerbi M. Speech processors for auditory prostheses. Bethesda (MD): National Institutes of Health, 1992. 11th Quarterly Progress Report, NIH project N01-DC-9-2401.
8. Pfingst BE. Operating ranges and intensity psychophysics for cochlear implants. Arch Otolaryngol 1984;110:140-4.
9. Shannon RV. Multichannel electrical stimulation of the auditory nerve in man. I. Basic psychophysics. Hear Res 1983;11:157-89.
10. Wilson BS, Finley CC, Lawson DT. Representations of speech features with cochlear implants. In: Miller JM, Spelman FA, editors. Cochlear implants: Models of the electrically stimulated ear. New York: Springer-Verlag, 1990:339-76.
11. Shannon RV. Psychophysics. In: Tyler RS, editor. Cochlear implants: audiological foundations. Chap. 3. San Diego (CA): Singular Publishing Group. In press.
12. White MW, Merzenich MM, Gardi JN. Multichannel cochlear implants: channel interactions and processor design. Arch Otolaryngol 1984;110:493-501.
13. Clark GM. The University of Melbourne-Nucleus multi-electrode cochlear implant. Adv Otorhinolaryngol 1987;38:1-189.
14. Wilson BS. Signal processing. In: Tyler RS, editor. Cochlear implants: audiological foundations. Chap. 2. San Diego (CA): Singular Publishing Group. In press.
15. Wilson BS, Lawson DT, Finley CC, Wolford RD. Coding strategies for multichannel cochlear prostheses. Am J Otol 1991;12(1 Suppl):56-61.
16. Rabiner LR, Shafer RW. Digital processing of speech signals. Englewood Cliffs (NJ): Prentice-Hall, 1978.
17. Bess FH, Townsend TH. Word discrimination for listeners with flat sensorineural hearing losses. J Speech Hear Disord 1977;42:232-7.
18. Dubno JR, Dirks DD. Evaluation of hearing-impaired listeners using a nonsense syllable test. I. Test reliability. J Speech Hear Res 1982;25:135-41.
19. Dowell RC, Brown AM, Mecklenburg DJ. Clinical assessment of implanted deaf adults. In: Clark GM, Tong YC, Patrick JF, editors. Cochlear prostheses. Edinburgh: Churchill Livingstone, 1990:193-205.
20. 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(128 Suppl):132-4.
21. 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;29:282-7.
22. Wilson BS, Lawson DT, Finley CC, Zerbi M. Speech processors for auditory prostheses. Bethesda (MD): National Institutes of Health, 1991. 9th Quarterly Progress Report, NIH project N01-DC-9-2401.