X

Original sent to printing

B

Third Quarterly Progress Report

November 1, 1989 through January 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. Analysis of data from studies with a recipient of the House Ear Institute (HEI) auditory brainstem implant.
2. Further analysis of data from studies with Symbion patient MP.
3. Presentation of project results at the Eleventh Annual Conference of the Engineering in Medicine and Biology Society (Finley, Nov. 8-12), the HEI (Finley, Nov. 15), and the Third Symposium on Cochlear Implants in Children (Wilson, Jan. 26 and 27).
4. Discussion with investigators at the HEI on the possibility of joint HEI/RTI development of highly-flexible speech processors for use outside of the laboratory by HEI and other implant patients.
5. Completion and acceptance of a paper on "Coding Strategies for Multichannel Cochlear Prostheses," to be published in the *Am. J. Otol.*
6. Continued preparation of other manuscripts for publication.

The primary purpose of this report is to present the results from activity 2 above. In particular, we show how consonant and vowel identifications are affected by parametric manipulations in the *continuous interleaved sampler, interleaved pulses*, and *peak-picker* processing strategies. These results complement those of our last progress report for this project, in that the last report provided a broad overview of findings from the best implementations of the first two strategies and the present report provides detailed results from all tested variations.

II. Evaluation of Alternative Implementations of the *Continuous Interleaved Sampler, Interleaved Pulses*, and *Peak-Picker* Processing Strategies

In an intensive series of studies with subject MP 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 (the CIS processor was referred to as the "supersampler" processor in QPR 2) were compared with each other and with MP's compressed analog (CA) processor using an a full battery of speech tests. Results from those comparisons are presented in QPR 2. The additional processor variations were evaluated with tests of consonant identification (all variations) and vowel identification (some variations). The purpose of this report is to present the phoneme identification results from all tested variations.

Processors

As mentioned in QPR 2, results from psychophysical studies with MP indicated that he could rank the six intracochlear electrodes in his Symbion implant in an appropriate tonotopic order. All processors tested with him made use of this ability by stimulating electrodes near the apex of the cochlea to indicate the presence of low-frequency sounds and by stimulating electrodes near the base of the cochlea to indicate the presence of high-frequency sounds. However, other details of the stimulation patterns were quite different among processing strategies and their variations.

Waveforms of the CA, IP 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. In particular, 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 (left column). In addition, the onset of the unvoiced /t/ burst is represented in the stimuli of the basal channels (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 associated with the use of CA processors is that many of the temporal details present in the 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.

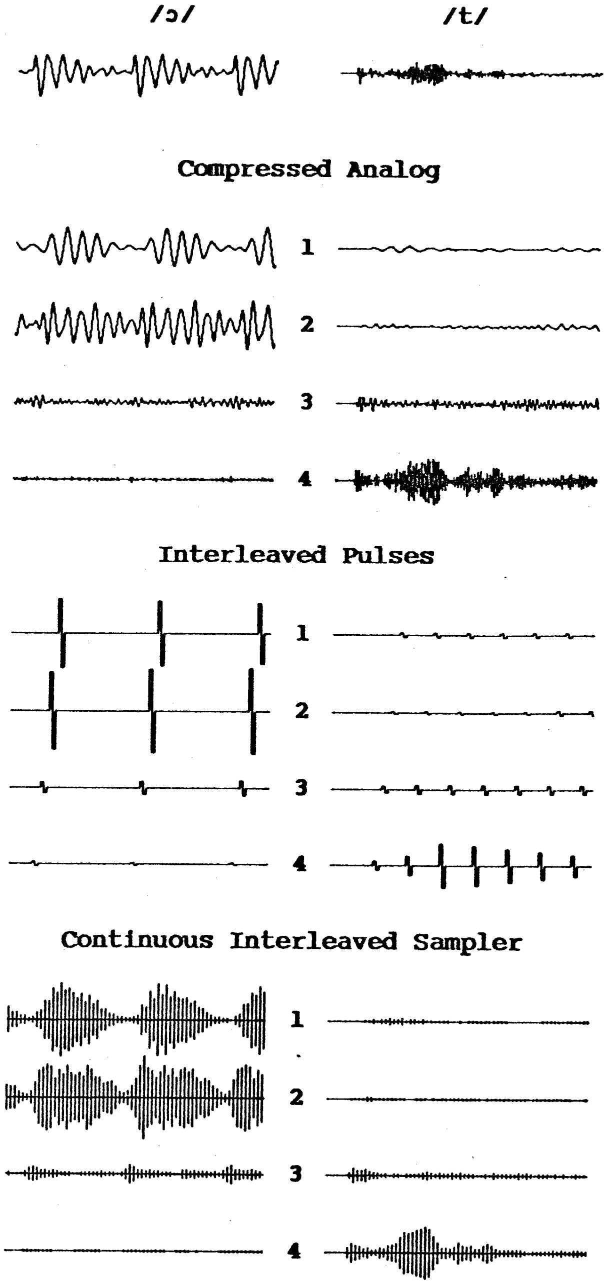

Fig. 1. Waveforms of three processing strategies. Equalized (6 dB/octave attenuation below 1200 Hz) speech inputs are shown at the top and stimulus waveforms for each of the strategies below. The left column shows input and stimulus waveforms for a voiced speech sound and the right column those for an 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 constants of the integrating filters for bandpass energy detection are 8.0 ms in the IP strategy and 0.4 ms in the CIS strategy. The duration of each trace is 25.4 ms.

The problem of channel interactions is addressed in the IP and CIS processors through the use of interleaved nonsimultaneous stimuli. There is no temporal overlap between stimulus pulses, so that 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 IP and CIS processors, pulse amplitudes are determined with a logarithmic transformation [Wilson et al., 1988a] of RMS energies to compress the dynamic range of those energies into the range of electrically-evoked hearing.

Differences between the IP and CIS processors include the rate of stimulation and the way in which voiced and unvoiced segments are treated. In the IP processor distinctions between voiced and unvoiced segments are represented by the timing of cycles of stimulation across the electrode array. During voiced segments stimulation cycles are presented at the fundamental frequency of the speech sound, and during unvoiced segments stimulation cycles are presented either at a fixed, high rate or at randomly-varied intervals.

In a known for what?

In contrast, the CIS processor presents stimulation cycles at the maximum rate (with one cycle immediately following its predecessor) during both voiced and unvoiced segments. In addition, this processor generally uses the shortest possible durations for pulses and intervals between pulses so that rapid variations in RMS energies can be followed by variations in pulse amplitudes for each channel.

Comparison of the three processors in Fig. 1 illustrates different tradeoffs between representations of temporal and spatial (channel-related) information. The CA processor provides a high degree of temporal detail, but also may have a high level of channel interactions. The IP processor presents a relatively sparse pattern of stimulation with concomitant loss of temporal detail. However, this loss may not be significant for most implant patients (see above), and, if significant, may be more than compensated for by reduction in channel interactions through the use of nonsimultaneous stimuli. Finally, temporal details are restored in the CIS processor through rapid stimulation rates on 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).

A fourth tradeoff between representations of temporal and spatial information is embodied in the PP processor. In this processor the position of a peak in either the bandpass output or RMS energy output of a channel is signaled by the presentation of a pulse. Also, as in the IP and CIS processors, the exact timing of the pulses is adjusted so that there is no temporal overlap of stimuli across channels.

The design of the PP processor is illustrated in Fig. 2. The upper panel shows bandpass outputs for each of four channels along with the stimulus pulses derived from those outputs. In addition, the positions of the peaks in the bandpass outputs are marked by short vertical lines above each trace. The lower panel shows the stimulus pulses only.

In the particular variation of PP processor shown in Fig. 2, a timer is advanced for each channel in a sequence of stimulation across the electrode array. At each time step a pulse is delivered if a peak occurred in the bandpass output between the previous and present time steps for

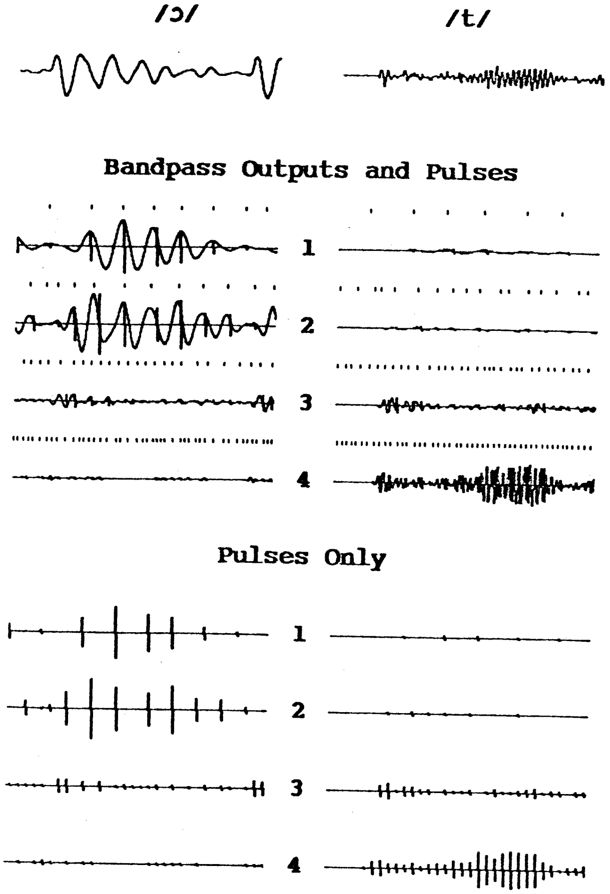

Fig. 2. Waveforms of the "peak picker" (PP) processing strategy. Equalized speech inputs are shown at the top and two panels of processor waveforms below. The upper panel shows the bandpass outputs and stimulus pulses for each of four channels. The locations of peaks in the bandpass outputs are marked with short vertical lines above each trace. The lower panel shows the stimulus pulses only. The duration of each trace is 12.25 ms.

that channel. The amplitude of the pulse is determined with the same logarithmic transformation used in the IP and CIS processors (i.e., the actual pulse amplitudes would be a logarithmic transformation of the amplitudes shown in Fig. 2). A fixed time is reserved for each channel in the stimulation sequence whether or not a pulse is delivered. As indicated in Fig. 2, this variation of the processor produces clusters of pulses at the F0 rate and individual pulses at the F1 rate for voiced speech sounds (left panels). Because the pulses must be presented nonsimultaneously, though, higher frequencies in the bandpass outputs are not followed with pulses at those frequencies. Notice, for instance, that many peaks are missed in channels 3 and 4, and that large offsets between the positions of peaks and subsequent pulses are seen in the waveforms of channel 2.

Alternative implementations of the PP processor are illustrated in Fig. 3. The top panel shows the waveforms of the implementation just described ("Bandpass Outputs, Time for Each Channel"); the middle panel an implementation in which the time allocated for each channel is *not* used if a pulse is not delivered ("Bandpass Outputs, Channels Skipped"); and the bottom panel an implementation in which RMS outputs are used instead of bandpass outputs ("RMS Outputs, Channels Skipped"). As might be expected, the synchronization of pulses to peaks is greatly improved when channels without a pending pulse are skipped in the stimulation sequence (middle panel).

A summary of waveforms for pulsatile processors is presented in Fig. 4. All three processors use nonsimultaneous stimuli. Among processors, the CIS processor provides the greatest density of temporal information and the IP processor the least. The PP processor provides an intermediate level of temporal detail, with a representation of F1 in the apical channel(s). In addition, the PP processor presents different rates of stimulation on each channel, which might increase the salience of channel-related cues for some patients.

Processor Parameters

Parameters of all processors used in the tests with MP are presented in Table 1. The variations of IP processors included manipulations of the time between sequential pulses and pulse polarity. Processor MP1L presented pulses with the positive phase leading (which produced a stimulus with an initially *cathodic* phase at the intracochlear electrode) and with a separation of 400 μs. Processor MP1LC was identical to MP1L except that the pulse polarity was reversed. Processor MP1L0 was identical to MP1L except that the time between sequential pulses was reduced to zero. Finally, processor MP1LS presented its pulses simultaneously across channels for each stimulation cycle.

All IP processors of the present study used six channels of stimulation, and all delivered stimulation sequences at the fixed rate of 278 Hz during unvoiced intervals. Other variations of IP processors were evaluated in a previous study with MP, and the interested reader is referred to QPR 1 of this project for the results.

The tested variations of CIS processors included manipulations of the frequency ranges spanned by the bandpass filters, the rectifier used in the RMS energy detector (full wave or half wave), and the update order of sequentially-stimulated channels. All CIS processors used six channels of stimulation. The rate of stimulation on each channel was 1515 Hz.

Only one implementation of the PP processor was tested with MP. This implementation used the design outlined in Fig. 2, i.e., with time taken for each channel whether or not a pulse is delivered. The PP processor used six channels of stimulation, with the same update order found to be best in the evaluations of the CIS processors (6-3-5-2-4-1).

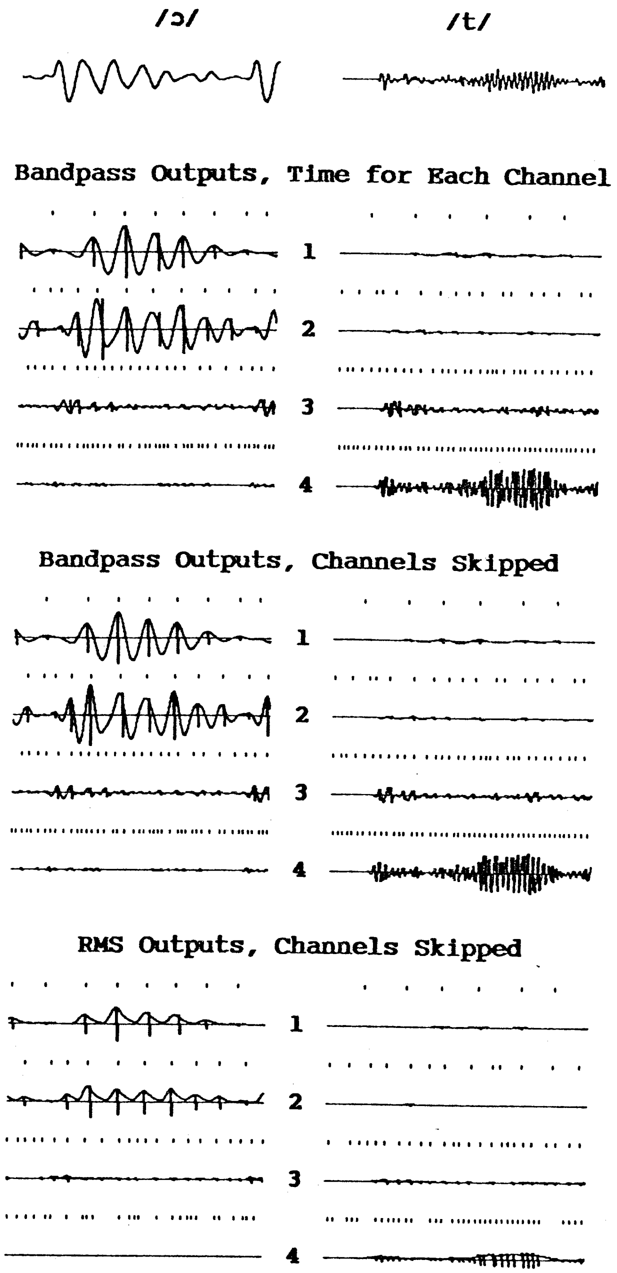

Fig. 3. Various implementations of the peak picker processor. Equalized speech inputs are shown at the top and waveforms for the processor implementations in the three panels below. The duration of each trace is 12.25 ms.

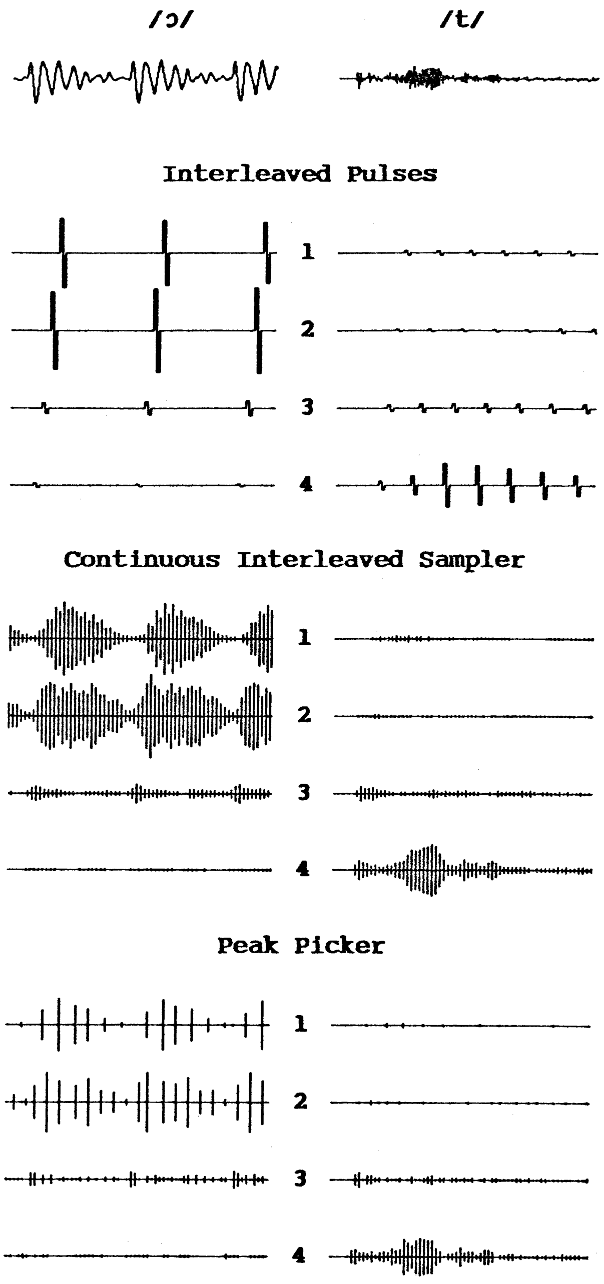

Fig. 4. Waveforms of three types of pulsatile processors. Equalized speech inputs are shown at the top with stimulus waveforms for each of the strategies below. The duration of each trace is 25.4 ms.

Table 2. Number of presentations of each token in the tests of consonant and vowel identification.

	Consonants		Vowels	
Processor	M	F	M	F
Symbion (CA)	25	25	9	6
MP1L	5			
MP1LS	10	10	9	9
MP1L0	10	10	9	6
MP1LC	5			
RTSS2	5			
RTSS3	5			
RTSS4	5			
RTSS8	10	10	9	9
RTPP1	10	10	6	6

Results

To evaluate the patterns of confusions (and correct responses) from the tests of consonant and vowel identification, matrices of responses for each processor were used as inputs to the information transmission (IT) analysis of Miller and Nicely [1955]. 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 subject. The consonant features selected for the present study were voicing (voice), nasality (nasal), frication (fric), duration (dur), place of articulation (place), and envelope cues (envel). The vowel features were first formant frequency (F1), second formant frequency (F2), and duration.

Results for the five processors evaluated with both consonant and vowel tests, and with male and female speakers, are presented in Figs. 5 and 6. The consonant results (Fig. 5) show clear differences among processors. In particular, comparison of scores for the CA (Symbion), IP (MP1L0) and CIS (RTSS8) processors demonstrate superior performance for the latter two processors despite the subject's four years of daily experience with his Symbion processor. Use of the IP processor produces gains over the CA processor for every feature except voicing (where scores are about the same), and use of the CIS processor produces gains over both the IP and CA

Table 1. Parameters of the interleaved pulses (MP1L$\underline{x}$), continuous interleaved sampler (RTSS$\underline{x}$), and peak-picker (RTPP1) processors. The pulse parameters include polarity (pol), duration per phase (dur/ph) and separation between pulses (sep), and the remaining parameters include the frequency range spanned by the bandpass filters (Bandpass range), the corner frequency of the integrating filters for bandpass energy detection (RMS integrator) and the frequency below which speech signals are attenuated for input equalization (eq). Entries for the last three measures are given in Hertz, and all remaining entries except polarity and channel update order are given in microseconds. Changes in parameters from one processor to the next are highlighted with **boldface type.**

IP + CIS processors

formance

Proc	Channel update order	pol	dur/ph µsec/ph	sep µsec	Time between pulse sequences µsec	Round-robin time µsec	Bandpass range	RMS integrator	Rect	eq
MP1L	6-5-4-3-2-1	+	100	400	400	3600	350-6500	25	FW	1200
worst → MP1LS	6-5-4-3-2-1	+	100	**N/A**	**3400**	3600	350-6500	25	FW	1200
MP1L0	N/A	+	100	**0**	**2400**	3600	350-6500	25	FW	1200
MP1LC	6-5-4-3-2-1	**-**	100	400	400	3600	350-6500	25	FW	1200
RTSS2	6-5-4-3-2-1	+	55	0	N/A	660	350-7000	800	FW	600
RTSS3	6-5-4-3-2-1	+	55	0	N/A	660	**350-6000**	800	FW	600
RTSS4	6-5-4-3-2-1	+	55	0	N/A	660	350-6000	800	**HW**	600
→ RTSS8	**6-3-5-2-4-1**	+	55	0	N/A	660	350-7000	800	FW	600
RTPP1	6-3-5-2-4-1	+	55	0	N/A	N/A	350-7000	N/A	N/A	600

CIS processing

~ 1500 Hz

Tests

Tests of consonant and vowel identification were used to evaluate the processors of Table 1. 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 and identification of 8 vowels (/i, ɔ, ɛ, u, I, U, ʌ, æ/) in a /h/-vowel-/d/ context. 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 of five randomized presentations of each consonant or three randomized presentations of each vowel for one of the speakers. The total number of presentations for the processors and tests of this study are presented in Table 2.

All tests were conducted with hearing alone, and without feedback as to correct or incorrect responses.

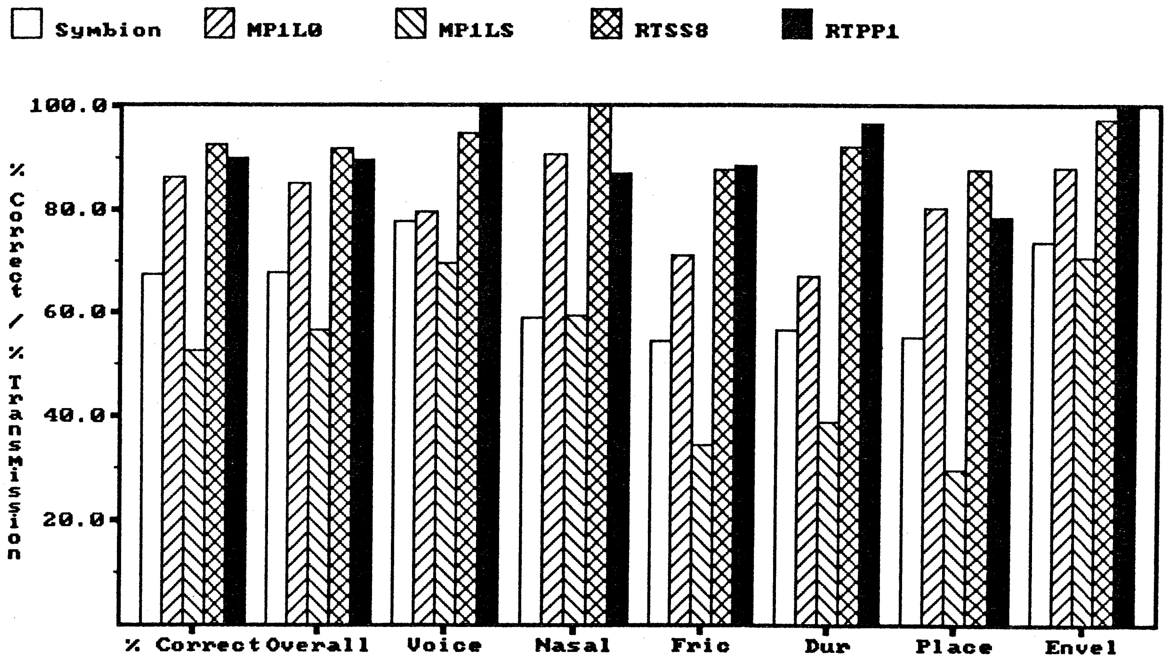

Fig. 5. Relative information transfer of consonant features for subject MP with various processors, hearing alone, combined male and female speakers. The features include voicing (voice), nasality (nasal), frication (fric), duration (dur), place of articulation (place), and envelope cues (envel).

processors for all features. In addition, substantial increases in the scores for voicing, frication, and duration are obtained when the CIS processor is used instead of the IP processor.

The remaining scores in Fig. 5 are those for the "simultaneous pulses" IP processor (MP1LS) and the PP processor (RTPP1). As might be expected from the previous discussion on channel interactions, use of the simultaneous pulses processor produces large decrements in the scores obtained with the otherwise identical IP processor (MP1L0). Especially large decreases are found for overall transmission, nasality, frication, duration, and place of articulation.

In contrast to the simultaneous pulses processor, the PP processor produced high IT scores. Indeed, the overall performance of the PP processor is roughly comparable to that of the CIS processor. Scores for voicing, duration, and envelope are all somewhat higher for the PP processor, and scores for nasality and place of articulation are lower for PP processor.

In comparison to the consonant results, the vowel scores (Fig. 6) are more uniform among the CA, IP and CIS processors. The three processors produce identical or nearly-identical scores for percent correct, overall transmission, F2, and duration. The CA processor produces a somewhat higher score for F1, however.

As before, application of the simultaneous pulses processor produces large decrements in performance. Most notable are the reductions in the scores for F1 and F2.

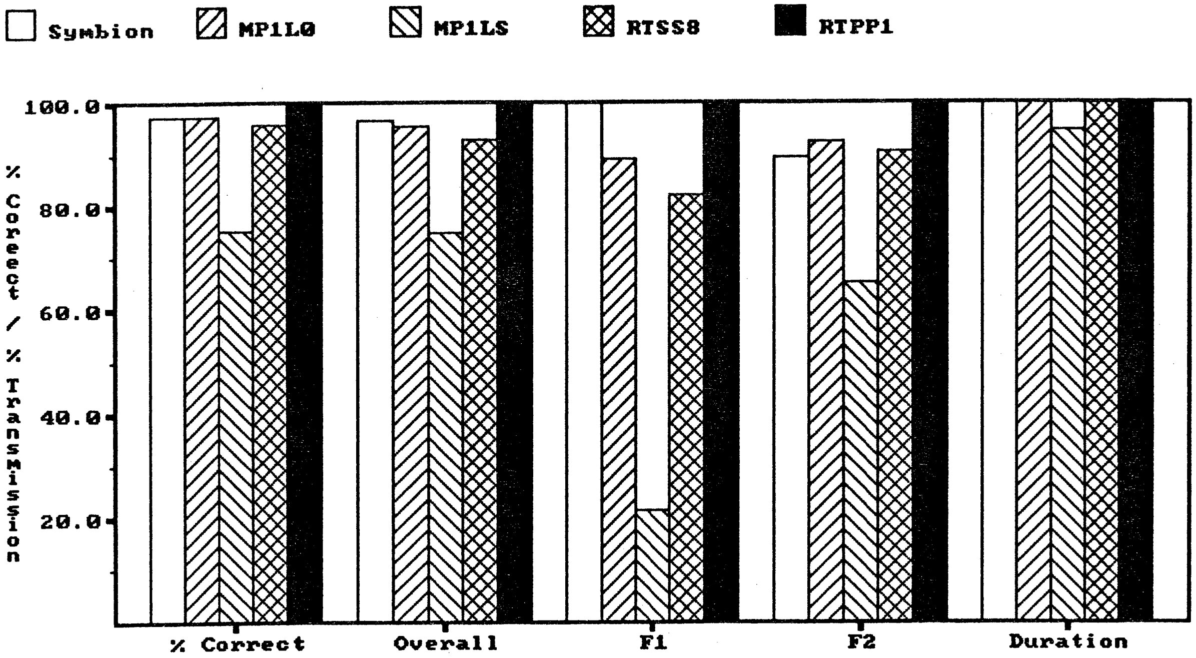

Fig. 6. Relative information transfer of vowel features for subject MP, for the same processors shown in Fig. 5, hearing alone, combined male and female speakers. The features include first formant (F1), second formant (F2), and duration.

The best performance on the vowel test was obtained with the PP processor. In fact, all presentations of the vowel tokens by both male and female speakers were perfectly identified by MP when using the PP processor.

Although performance with the PP processor was perfect for the vowel test, it should be noted that performance was nearly perfect with the CA, IP and CIS processors. Thus, ceiling effects might have masked differences that may exist among these processors.

The remaining processors listed in Tables 1 and 2 were briefly evaluated with the consonant test using the male speaker. The results from that test are presented in Fig. 7 for all tested variations of IP processors (including MP1L0 and MP1LS) and in Fig. 8 for all tested variations of CIS processors (including RTSS8). As indicated in Fig. 7, three of the four variations of IP processors produced similar percent correct and overall transmission scores. These three variations (MP1L, MP1L0 and MP1LC) all used nonsimultaneous pulses. The variation with simultaneous stimulation across channels (MP1LS) produced markedly lower percent correct and overall transmission scores. In addition, the simultaneous pulses processor produced much lower scores than the other three processors for the features of duration and place of articulation.

In examining the results for the three best IP processors, it appears that changing the pulse polarity from positive leading (MP1L) to negative leading (MP1LC) improves the transmission of

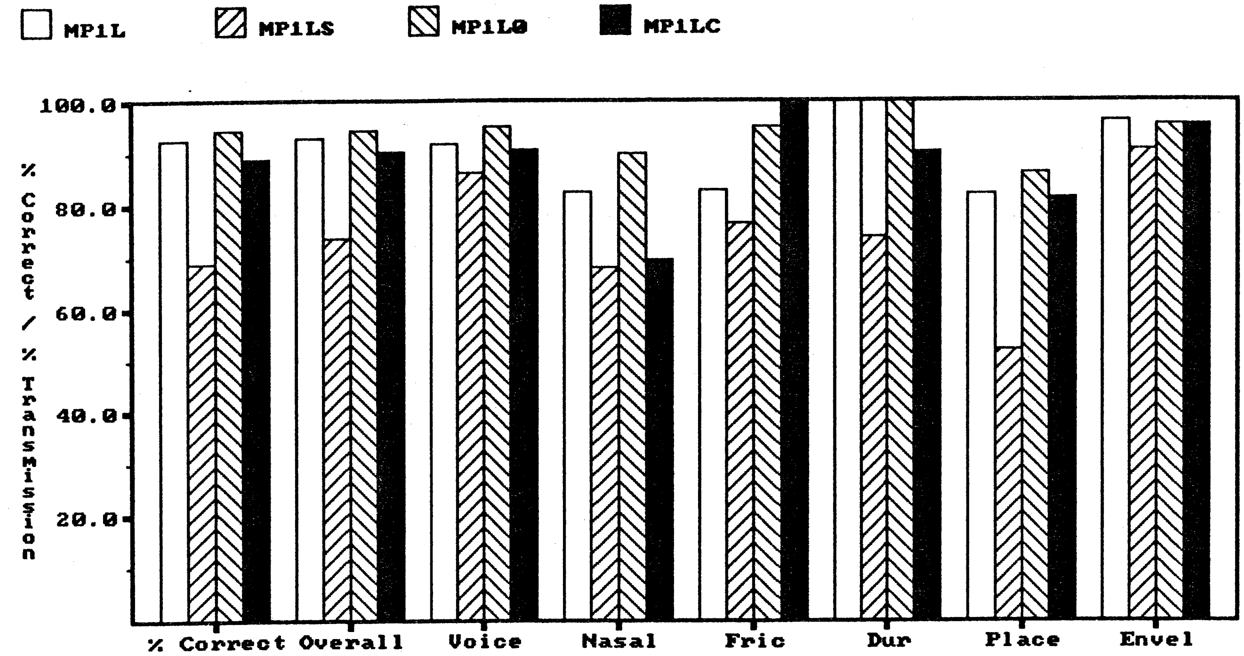

Fig. 7. Relative information transfer of consonant features for subject MP; various IP processors, hearing alone, male speaker.

fricative information, but degrades the transmission of nasality and duration information. In addition, reducing the time between pulses to zero (MP1L0) produces slight increases in the scores for voicing, nasality, and place of articulation.

Finally, all tested variations of CIS processors (Fig. 8) produced similar results. The only clear difference among these processors was a small decrement in the transmission of duration information for the processor with the half-wave rectifier in its RMS energy detectors (RTSS4). All four CIS processors produced perfect transmission scores for voicing, nasality, frication, and envelope cues.

Discussion

Three types of pulsatile processors were compared with each other and with the CA processor of the Symbion device using tests of consonant and vowel identification. With the exception of the "simultaneous pulses" IP processor, all types and variations of the pulsatile processors produced large gains over the CA processor in the transmission of consonant information. In particular, the IP, CIS and PP processors all produced increases in the IT scores for every consonant feature. Especially large increases shared by all three pulsatile processors included those for nasality and place of articulation. In addition, the CIS and PP processors produced

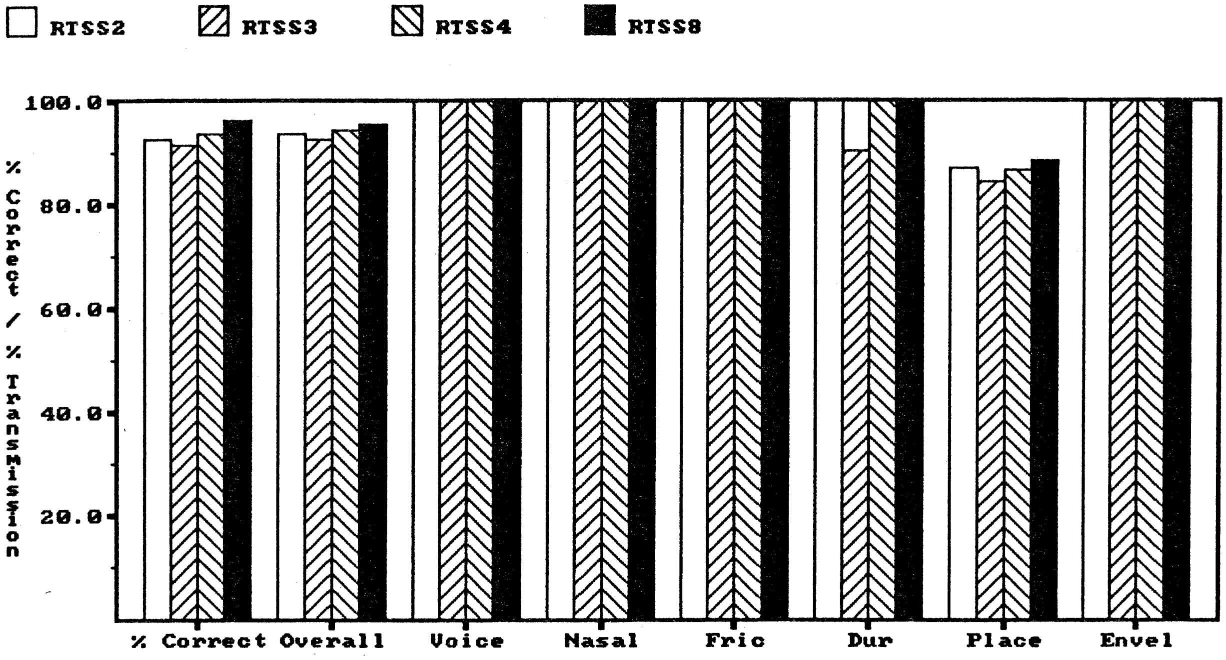

Fig. 8. Relative information transfer of consonant features for subject MP; various CIS processors, hearing alone, male speaker.

substantially higher scores for voicing, frication, duration, and envelope than did the IP processor.

These results may be understood in terms of processor design. As discussed above, the use of nonsimultaneous stimuli eliminates a principal component of channel interactions. This, in turn, might be expected to increase the salience of channel-related cues and to increase the transmission of consonant features differentiated by those cues. Two such features are nasality and place [Miller and Nicely, 1955; Wilson et al., 1990], and IT scores for these features are improved by each of the pulsatile processors using nonsimultaneous stimuli.

An additional aspect of processor design is the representation of temporal events. Both the CIS and PP processors present more temporal information than the IP processor, and both produce higher IT scores for the remaining consonant features, which are largely or solely represented by temporal variations in the speech signal [Miller and Nicely, 1955; Van Tassell et al., 1987].

The relatively poor showing of the CA processor in the transmission of temporal information may have been a consequence of severe channel interactions. Because the phase relationships among channels of stimulation are not controlled in the CA processor, a continuously changing (and unpredictable) pattern of channel interaction is produced during ongoing speech. It is highly likely that these changing patterns distort the representation of temporal cues in the neural responses to CA stimuli.

In addition to the problem of continuously varying channel interactions, the CA processor also presents much higher frequencies of stimulation in the basal channels than any of the tested pulsatile processors. Both psychophysical [Shannon, 1983] and single-unit [Moxon, 1971; Parkins, 1986; van den Honert and Stypulkowski, 1987; Javel, 1990] studies have demonstrated strong adaptation to high-frequency stimuli. Thus, use of such stimuli in the CA processor could produce temporal distortions in the representation of sustained high-frequency components in speech, a problem that may be avoided with the relatively-low frequencies of stimulation used in the pulsatile processors. A faithful representation of sustained high-frequency components would allow discrimination of the long-duration consonants (/ʃ, s, z/) from the remaining consonants in our test. This discrimination would improve the IT score for the duration feature. Also, the IT scores for envelope cues and frication might be enhanced by a representation that maintained perception of sustained high-frequency sounds.

The fact that high frequencies of stimulation can produce adaptation in the nerve suggests a reason for caution in the general application of CIS processors. That is, even though the frequencies used in the CIS processor are substantially lower than the maximum frequencies of the CA processor, the CIS frequencies may be high enough to produce at least some adaptation in some patients. In our experience, patients who exhibit various presumptive signs of poor nerve survival, such as high thresholds, narrow dynamic ranges, and high channel interactions, also exhibit loudness adaptation to sustained high-frequency stimuli. Therefore, these patients might be best served with the IP processor, which presents pulses at rates that do not produce loudness adaptation (e.g., rates below 300 Hz).

An intermediate approach with respect to adaptation would be to use the PP processor. Its rates of stimulation are generally lower than those of the CIS processor and higher than those of the IP processor. Thus, the PP processor may provide temporal details while still not producing loudness adaptation. The findings to date of improvements in the transmission of duration, envelope, and voicing features with the PP processor support this concept.

Finally, improved transmission of vowel features with the PP processor is consistent with an improved temporal representation in the apical channels. In particular, pulses are presented at the rate of F1 in those channels during voiced speech sounds. The large increase in F1 transmission obtained with the PP processor (over that obtained with the IP and CIS processors) may reflect perception of frequency changes in the apical channels.

Future Directions

As indicated in the Discussion, different processing strategies may be best for different patients. A major goal of future studies will be to evaluate the CIS and PP processors across a broad population of subjects, and to continue our comparative studies of the CA and IP processors with those same subjects.

Acknowledgements

We are pleased to acknowledge the collaboration of R.D. Wolford, D.K. Eddington and W.M. Rabinowitz in the studies with MP. We also are indebted to MP for his precise descriptions of speech percepts and for his enthusiastic participation.

References

Javel E: Acoustic and electrical encoding of temporal information. In Miller JM, Spelman FA (eds): *Cochlear Implants: Models of the Electrically Stimulated Ear*. New York: Springer-Verlag, 1990, pp 247-296.

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.

Moxon EC: Neural and mechanical responses to electric stimulation of the cat's inner ear. Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1971.

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

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

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

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

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

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.

III. Plans for the Next Quarter

Our plans for the next quarter include the following:

1. Presentation of project results in invited lectures 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).
2. Continued preparation of manuscripts for publication.
3. 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 (studies with four subjects are scheduled for the next quarter).
4. 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).
5. 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.
6. Continued development of portable processors based on the TMS320C25 device.

Appendix 1

Summary of Reporting Activity for the Period of

November 1, 1989 through January 31, 1990

NIH Contract N01-DC-9-2401

The following publications and presentations were made in the last quarter of project work. The paper for the IEEE/EMBS Meeting is reproduced on the next two pages.

I. Publications

Wilson, B.S., D.T. Lawson, C.C. Finley and R.D. Wolford: Coding strategies for multichannel cochlear prostheses. Accepted for publication in the *Am.. J. Otol.*

Finley, C.C.: A finite-element model of radial bipolar field patterns in the electrically stimulated cochlea -- Two and three dimensional approximations and tissue parameter sensitivities. *Proc. IEEE Engineering in Medicine and Biology Society 11th Annual Conf.*, New York: IEEE Press (CH2770-6/89/0000-1059), 1989, pp. 1059-1060.

II. Presentations

Finley, C.C.: Chairman of the session on cochlear prostheses, Eleventh Ann. Conf. Engineering in Medicine and Biology Soc., Seattle, WA, November 8-12, 1989.

Finley, C.C.: A finite-element model of radial bipolar field patterns in the electrically stimulated cochlea -- Two and three dimensional approximations and tissue parameter sensitivities. Presented at the Eleventh Ann. Conf. Engineering in Medicine and Biology Soc., Seattle, WA, November 8-12, 1989.

Finley, C.C.: Invited speaker presentation, House Ear Institute, Los Angeles, November 15, 1989.

Wilson, B.S.: Processing strategies for cochlear implants. Invited presentation at the Third Symposium on Cochlear Implants in Children, Indiana University School of Medicine, Indianapolis, Indiana, January 26 and 27, 1990.

A FINITE-ELEMENT MODEL OF RADIAL BIPOLAR FIELD PATTERNS IN THE ELECTRICALLY STIMULATED COCHLEA - TWO AND THREE DIMENSIONAL APPROXIMATIONS AND TISSUE PARAMETER SENSITIVITIES

Charles C. Finley

Neuroscience Program, Research Triangle Institute, Research Triangle Park, NC 27709
Div. of Otolaryngology, Dept. of Surgery, Duke University Medical School, Durham, NC

ABSTRACT

Field patterns within the electrically stimulated cochlea are described by a three-dimensional finite-element model. A radially-oriented bipolar electrode pair in an insulating carrier is modeled in scala tympani. Results are compared with an earlier two-dimensional model and field pattern sensitivities for selected tissue resistivities are explored.

INTRODUCTION

Previous work using finite-element models of intracochlear field patterns relied on a two-dimensional sheet description of a cochlear cross section with a bipolar electrode in scala tympani [1]. This earlier model had two distinct limitations. First, the aspect ratios for many triangular elements were not optimal. Second, the two-dimensional geometry made the implicit assumption that all regions of the cochlear cross section extended to infinity both above and below the plane of the cross section. This was a useful approximation for the soft tissue structures but was a significant distortion of the electrode surfaces. This paper presents estimates of field patterns produced by a more realistic three-dimensional, finite-element model of a cochlea implanted with a radially-oriented, bipolar electrode pair. Results from this model using other electrode configurations have been published previously [2].

THE THREE-DIMENSIONAL FINITE-ELEMENT MODEL

In this model a cross section of the cochlea is projected linearly along an axis perpendicular to the plane of the section, thus producing a short, straight segment of the cochlea. A radially-oriented bipolar electrode pair, mounted in a carrier insulator is located in scala tympani. An enlarged view of the central region of a layer from the model is shown in Figure 1. Each finite element is shrunk 15% geometrically for emphasis. Each layer of the model contains 304 nodes defining 204 five- and six-sided solid elements. Twelve layers of varying thicknesses are included in the complete model for a total of 1976 nodes and 2448 solid elements. Results for the two electrode configurations shown in Figure 2 are presented in this paper. One configuration (a) assumes the electrodes extend the full length (5.2mm) of the longitudinal axis of the model, whereas the other configuration (b) is a discrete focal pair spanning only 200 microns of the longitudinal axis. An electrode configuration is specified by setting the resistivities of the surface plate elements of the electrode carrier to represent either insulators or electrode conductors. Node potentials at nodes bordering the conductive electrode surfaces are fixed arbitrarily at +100 and -100 mvolts prior to computation. Element resistivities (in ohm-cm) are defined regionally to characterize the electrodes(0.1), the carrier insulator(10^9), the endolymph(60), the perilymph(70), Reisner's membrane(60480), basilar membrane(1800), the anisotropic neural tissue of the peripheral axon leading down from the habenula to the spiral ganglion(300 axial; 1500 transaxial), the spiral ganglion itself(300) and bone(630). The model sensitivity to tissue resistivity specifications is studied by

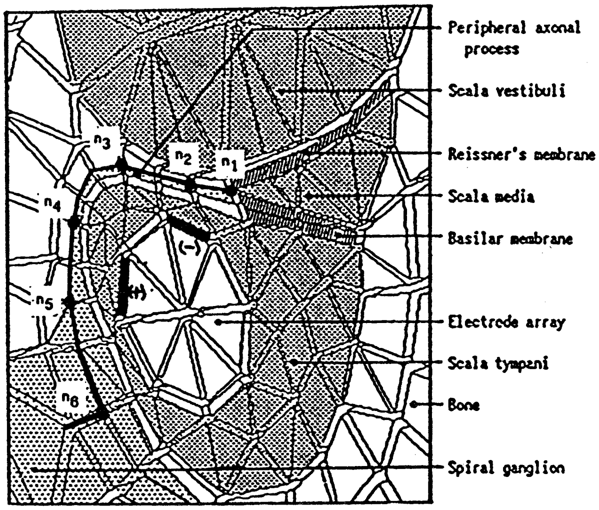

Figure 1. Single layer of FE model.

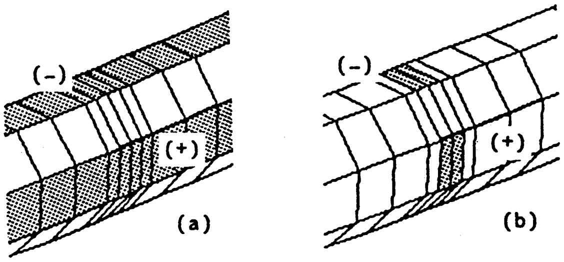

Figure 2. Electrode configurations (see text).

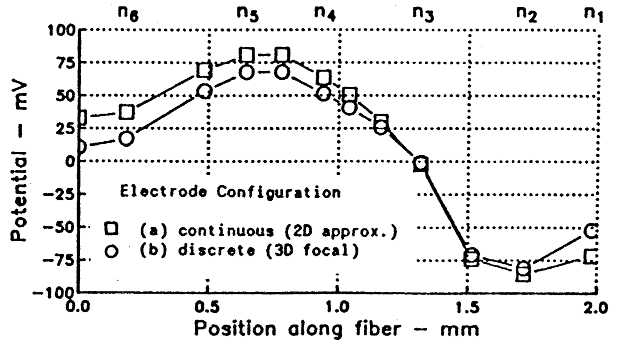

Figure 3. Field patterns for electrodes (a) and (b).

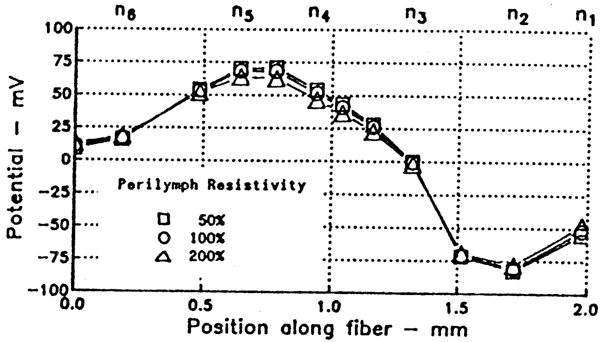

Figure 4. Perilymph effects on field patterns.

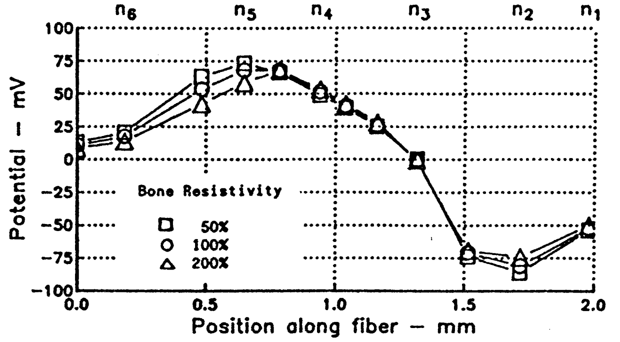

Figure 5. Bone effects on field patterns.

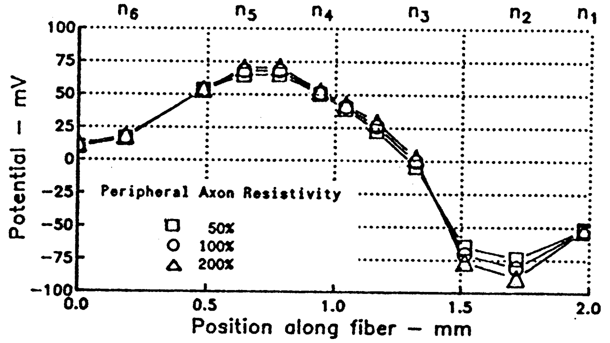

Figure 6. Peripheral axonal process effects on field patterns.

independently setting the resistivities of perilymph, bone and the peripheral axon to 50% and 200% of their standard values listed above. The outputs of the model are the potential patterns that occur in the vicinity of the neural elements for fixed potentials applied to the electrodes. For this paper, potentials along a single neural locus, located directly above the center of the electrode structure, are reported. This locus lies on the modiolar side of the peripheral axonal processes and extends from the center of the ganglion to the habenula as indicated in Figure 1. Further details of the model construction can be found in [2].

RESULTS

Figure 3 shows the neural element potential levels for both electrodes described in Figure 2. Note that the electrode configuration that approximates the implicit assumptions of the previous two-dimensional model produces larger peak levels and steeper gradients in the region between the electrode pairs. However, at distant neural locus positions, the spatial gradients are significantly smaller.

Figures 4, 5 and 6 show the effects on neural locus potential levels for 50% and 200% manipulations of perilymph, bone and peripheral axonal process resistivities, respectively. In all cases, alterations of the potential levels are modest and the lower resistivity condition produces a more broadly spreading electrode field. Perilymph resistivity mostly affects potential levels between the electrodes, since the largest volume of perilymph in the vicinity of the electrode lies between the electrodes. Bone resistivity mostly affects potential levels in the ganglion region, due to the greater bone mass located near the ganglion in the modiolus. Only very thin bony elements surround the peripheral neural processes. Neural tissue resistivity has a slight impact on field patterns near the habenula in this model.

CONCLUSIONS

Based on this more realistic three-dimensional model, the conclusions from previous work [1] are confirmed and restated. For a closely-placed, radially-oriented, bipolar scala tympani electrode:

- o first order effects on potential patterns are due to electrode geometry;
- o second order effects on potential patterns are due to tissue impedance effects.

REFERENCES

1. Finley C, Wilson B and White M (1987): A finite-element model of bipolar field patterns in the electrically stimulated cochlea: A two-dimensional approximation. IEEE/9th Conf. EMBS, 1901-1903.

2. Finley C, Wilson B and White M (1989): Models of neural responsiveness to electrical stimulation. To be published as chapter 5 in Models of the Electrically Stimulated Cochlea, (Miller JM and Spelman FA, eds.), Springer-Verlag.

ACKNOWLEDGMENT

This work was supported by contracts N01-NS-3-2356 and N01-NS-5-2396 from the Neural Prosthesis Program of NINCDS and a Professional Development Award from Research Triangle Institute.