<u>Seventh Quarterly Progress Report</u>

October 1 through December 31, 2003
NIH Project N01-DC-2-1002

Speech Processors for Auditory Prostheses

Prepared by

Blake Wilson, Robert Wolford,

Reinhold Schatzer, Xiaoan Sun and Dewey Lawson

Center for Auditory Prosthesis Research
Research Triangle Institute
Research Triangle Park, NC

CONTENTS

I. Introduction

The main objective of this project is to design, develop, and evaluate speech processors for implantable auditory prostheses. Ideally, such processors will represent the information content of speech in a way that can be perceived and utilized by implant patients. An additional objective is to record responses of the auditory nerve to a variety of electrical stimuli in studies with patients. Results from such recordings can provide important information on the physiological function of the nerve, on an electrode-by-electrode basis, and can be used to evaluate the ability of speech processing strategies to produce desired spatial or temporal patterns of neural activity.

Work and activities in this quarter included:

- Completion of studies with Subject ME19 through October 3, to evaluate combined electric and acoustic stimulation of the auditory system, for patients having some residual, low-frequency hearing.
- Continuing studies throughout the quarter with local subjects ME14, ME16 and ME22, implanted bilaterally and using Med-El Tempo+ devices. The studies with ME16 included use of new melody identification assessment tests.
- Attendance by Blake Wilson, with Duke University Medical Center surgeon Debra Tucci, at meetings with cochlear implant clinicians and researchers in Warsaw, Poland, October 14-15.
- Organization and conduct of a one-day workshop on “Future directions for the further development of cochlear implants,” by Rainer Klinke, Rainer Hartmann, and Blake Wilson, held in Frankfurt, Germany, October 15 and attended by approximately 25 experts in the field.
- Recognition as Guest of Honor and invited presentation by Blake Wilson at the Hearing Preservation Workshop, Frankfurt, Germany October 16-19.
- Attendance by Blake Wilson, Dewey Lawson, Reinhold Schatzer, and Xiaoan Sun at the Neural Prosthesis Workshop, NIH, October 21-23.
- A presentation by Dewey Lawson and Reinhold Schatzer at the Workshop.
- A visit by Chris Turner, University of Iowa, October 24.
- First studies with Nucleus Contour Electrode percutaneous subject NP6, December 1-2.
- A visit by consultant Marian Zerbi December 5-7 for verification of new processor code before its first use with a subject
- Two weeks of studies with return Ineraid percutaneous subject SR3, December 8-19, for evaluations of new dual-resonance nonlinear (DRNL) filter processing strategies, some including the use of DRNL filters in conjunction with virtual channels.
- Evaluation of additional strategies with this subject, designed to represent “fine structure” or “fine frequency” information with cochlear implants.
- Implantation surgeries for the third and fourth Nucleus Contour Electrode percutaneous study patients (NP7 and NP9) at Duke University Medical Center, Durham, NC., October 8 and December 10.
- Completion of a stereophonic version of complex tone synthesis software, for research studies with EAS subjects and subjects with binaural implants.

In the present report, we describe a new approach, that combines DRNL filters with virtual-channel processing. As noted above, this approach and others were evaluated in studies with Ineraid subject SR3. Results from studies conducted during the past quarter with other subjects, and from additional studies with SR3, will be reported in future reports.

II. Combined use of DRNL filters and virtual channels

In the Sixth Quarterly Progress Report for this project, we described an approach toward a closer mimicking with cochlear implants of the signal processing steps that occur in the normal inner ear. The approach included the use of "dual-resonance nonlinear" (DRNL) filters to reproduce the nonlinear filtering properties of tuning and compression in the normal inner ear, that result from a feedback system that includes the basilar membrane and outer hair cells (Meddis et al., 2001; Lopez-Poveda and Meddis, 2001; Wilson et al., 2003). Details of the DRNL filters and their incorporation into designs of speech processors for cochlear implants are presented in QPR 6 (Schatzer et al., 2003).

One finding from the initial studies reported in QPR 6 is that the DRNL filters are much more sharply tuned than the standard Butterworth filters used in conventional designs, especially at low input levels and especially for a relatively small number of stimulus sites and associated channels. The properties of the DRNL filters of course reflect the sharp tuning found in the normal inner ear, especially at low sound pressure levels (SPLs).

Preliminary results were presented in QPR 6 for seven users of bilateral Med-El implants. The maximum number of available stimulus sites on each side was 8-12, depending on the subject, side of the implant, and implant type (COMBI 40 versus COMBI 40+). This small number of sites produced a "picket fence" effect when DRNL filters were applied. In particular, and especially at low input levels, gaps were produced in the represented spectra of input sounds due to the relatively sharp tuning of the filters. This may have limited the performance of processors using one DRNL filter per stimulus site.

In two variations of processors using DRNL filters, we assigned more than one filter to each stimulus site. In one variation, the average of outputs from multiple DRNL channels was calculated and then that average was used to determine the amplitude of a stimulus pulse for a particular electrode. Each DRNL channel included the DRNL filter, an envelope detector, and a lookup table for compressive mapping of envelope levels onto pulse amplitudes. Thus, the average was the average of mapped amplitudes for the number of DRNL channels assigned to the electrode. We called this the "avg *n*-to-*m* approach," in which *m* was the maximum number of electrodes available in the implant and in which *n* was the total number of DRNL channels, an integer multiple of *m*. In another variation, the maximum of outputs from the channels for each electrode was identified and then that maximum was used to determine the amplitude of the stimulus pulse. We called this the "max *n*-to-*m* approach."

These *n*-to-*m* approaches are illustrated in Fig. 1. As shown, the spectral gaps produced by assigning only one DRNL filter to each stimulus site (thin lines, top panel) is reduced or largely eliminated with either the avg *n*-to-*m* or max *n*-to-*m* approaches.

As noted in QPR 6, use of these approaches produced significant increases in speech reception scores in some cases, compared with processors that simply assigned the output

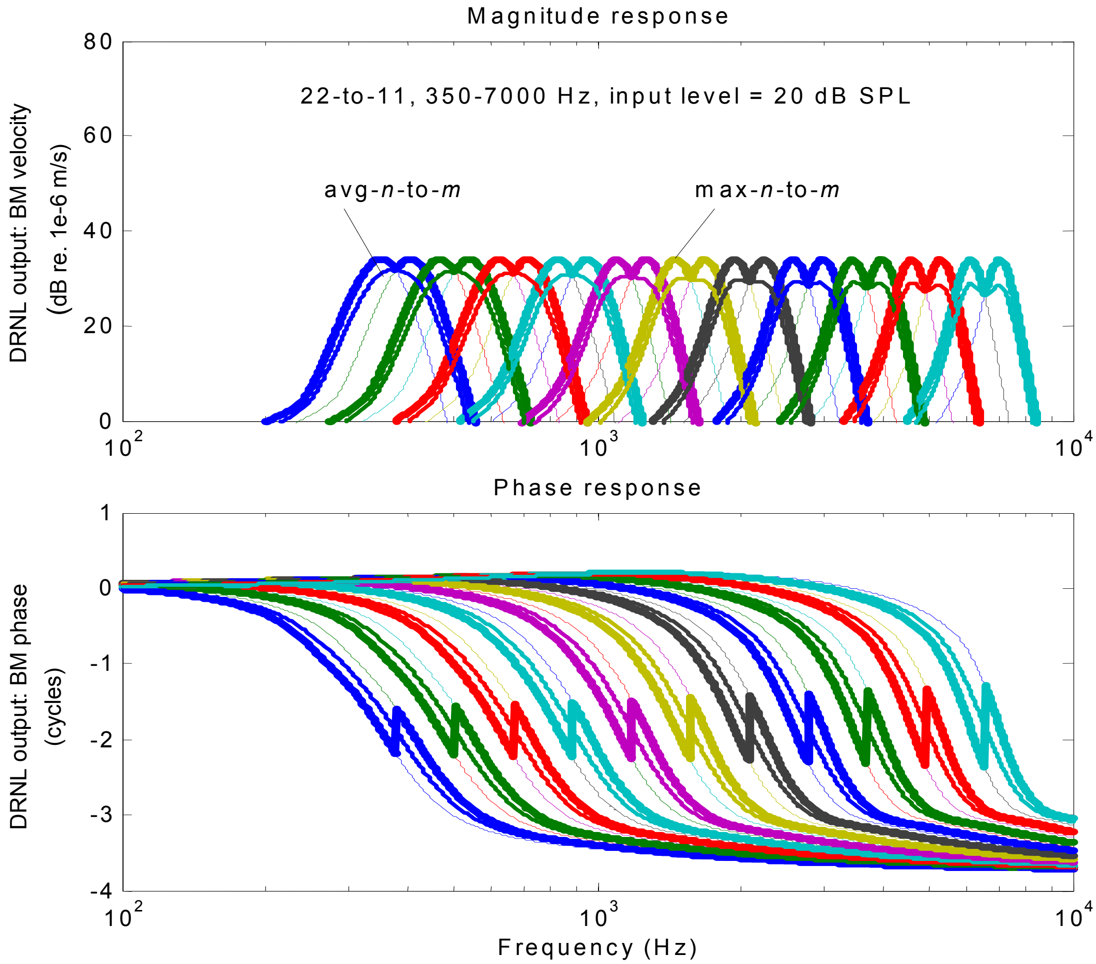

Figure 1. The *n*-to-*m* approach for avoiding spectral gaps in a processor design based on a DRNL filter bank: Let *m* be the number of available electrodes, then $n = k*m$ is the number of DRNL filters in the filter bank, with *k* an integer number greater than 1. The compressed envelope signals from *k* adjacent channels are combined, using either the average or selecting the maximum, to modulate the carrier pulse trains on each of the *m* electrodes. The particular curves shown illustrate the combined magnitude and phase response for a *22*-to-*11* DRNL filter bank at a low input level, where the filters are narrowest. The thin lines represent the output from the 22 individual DRNL filters. The medium-thick lines represent the *22*-to-*11* approach, using the average. The thickest lines represent the *22*-to-*11* approach, using the maximum. Note that the combined magnitude and phase response here is illustrated for DRNL filter band output signals, whereas in a speech processor implementation, the compressed envelope signals of neighboring filter outputs are combined, as opposed to the band outputs directly. As a consequence, the effect of the envelope extraction and maplaw compression stage is not accounted for here. (This figure and caption are reproduced from QPR 6, Schatzer et al., 2003.)

of each DRNL channel to a single corresponding stimulus site. Improvements for speech reception in noise were generally larger than improvements for speech reception in quiet. In a few cases where comparisons were made, the max *n*-to-*m* approach was better than the avg *n*-to-*m* approach. The best of the DRNL processors using an *n*-to-*m* approach produced speech reception scores that were as good as, but not better than, control CIS processors using *m* channels and *m* sites of stimulation.

We regarded this as an encouraging result, an immediate matching of performance with a new processing strategy, with very little or no experience in using the new strategy. In many prior studies, we and others (e.g., Tyler et al., 1986) have found that such an initial equivalence can be followed by much better performance with the new strategy, once subjects gain some experience with the new strategy.

At the same time, we recognized several possibilities for improvement in the design of processors using DRNL filters, that might produce even higher levels of initial performance. Those possibilities included (1) further adjustment and testing of the many parameter values in DRNL filters; (2) combination or selection of DRNL filter outputs, rather than the DRNL channel outputs, in designs using *n*-to-*m* approaches; and (3) using the same number of filters as stimulus sites, but with a high number of stimulus sites.

The first of these possibilities recognizes that the parametric space within and across DRNL filters is quite large. We have just begun to explore this space.

The second possibility recognizes that considerable distortions and complexities may be produced in combining or selecting signals that have been altered by a highly nonlinear mapping function, in addition to the nonlinearities of the DRNL filters. Combination or selection of the filter outputs, prior to envelope detection and (further) nonlinear processing, might be better than combination or selection following all of these operations.

The third possibility might retain the likely advantages of DRNL filters (that may result from compression and nonlinear tuning, as in normal hearing) while not discarding or distorting information as is inherent in the *n*-to-*m* approaches. The spectral gap problem would be handled through the use of a high number of stimulus sites, rather than with one of the *n*-to-*m* approaches.

In the present studies, we compared three basic processor designs in tests with a user of the Ineraid device, which includes a percutaneous connector and six intracochlear electrodes. The designs included a processor using 24 DRNL channels mapped to the six electrodes using a max *24*-to-*6* approach, as described above and in greater detail in QPR 6. The parameter choices used for the DRNL filters included a set to provide a flat frequency response across the spectrum spanned by all the filters, as also described in QPR 6. The spectrum was from 350 to 7000 Hz. This processor is referenced in the remainder of this report as the “cp CIS” processor. (The characters “cp” refer to a DRNL filterbank that is designed to provide a close replication of responses to sound at the *cochlear partition*.) Other aspects of the processor, such as the interlacing of stimuli across electrodes, are the same as in the standard CIS strategy.

The other two processor designs employed virtual channels as a way to increase the number of discriminable stimulus sites beyond the number of actual electrodes. This concept was introduced by our team in the early nineties (Wilson et al., 1992, 1993, 1994, 1995), and has since been investigated by others (Litvak et al., 2003; Poroy and Loizou, 2001). In the report by Litvak et al., the term "current steering" is used instead of the term "virtual channels," to reference the same concept.

A series of diagrams illustrating the construction of virtual channels is presented in Fig. 2. With virtual channels (or current steering), adjacent electrodes may be stimulated simultaneously to shift the perceived pitch in any direction with respect to the percepts elicited with stimulation of one of the electrodes only. Studies with implant subject SR2 and many others indicate that pitch can be manipulated through various choices of simultaneous and single-electrode conditions. If, for instance, the apicalmost electrode of the Ineraid array (electrode 1) is stimulated alone, SR2 and other subjects have reported a low pitch. If the next electrode in the array (electrode 2) is stimulated alone, a higher pitch is reported. An intermediate pitch can be produced for all subjects studied to date by stimulating the two electrodes together with identical, in-phase pulses. Finally, by reversing the phase of one of the simultaneous pulses, pitch percepts higher or lower than those produced by stimulation of either electrode alone can be produced. For example, a pitch lower than that elicited by stimulation of electrode 1 only can be produced by simultaneous presentation of a (generally smaller) pulse of opposite polarity at electrode 2. The availability of pitches other than those elicited with stimulation of single electrodes only may provide additional discriminable sites along (and beyond) the length of the electrode array. Such additional sites may support additional, perceptually separable, channels of stimulation and reception. We call these additional channels "virtual channels," and processors that use them *virtual channel interleaved sampling* (VCIS) processors.

The two additional processor designs included in the comparisons of the present studies used a VCIS approach to provide 21 discriminable sites of stimulation with Ineraid subject SR3's array of six intracochlear electrodes. The approach is illustrated in Fig. 3, in which stimulus site 1 is produced by stimulation of electrode 1 only, stimulus site 2 by simultaneous stimulation of electrodes 1 and 2 with a pulse amplitude of 75 percent for electrode 1 and of 25 percent for electrode 2, and so on. As indicated in the caption to the figure, results from pitch-ranking tests, using a 2AFC procedure, indicated that each of the 21 sites thus formed produced a distinct pitch, i.e., a pitch that is significantly different from those produced by stimulation of the neighboring site(s).

We note that even this fine resolution may not fully exploit SR3's perceptual abilities. In pilot studies, we also evaluated pitch ranking with 10 percent steps in current ratios for electrodes 1 and 2, and for electrodes 5 and 6, as opposed to the 25 percent steps used in the ranking tests of Fig. 3. SR3 was able to rank these closely spaced sites, with the 10 percent changes in current ratios, 4 out of 4 times for all pairings. (The tests with the 25

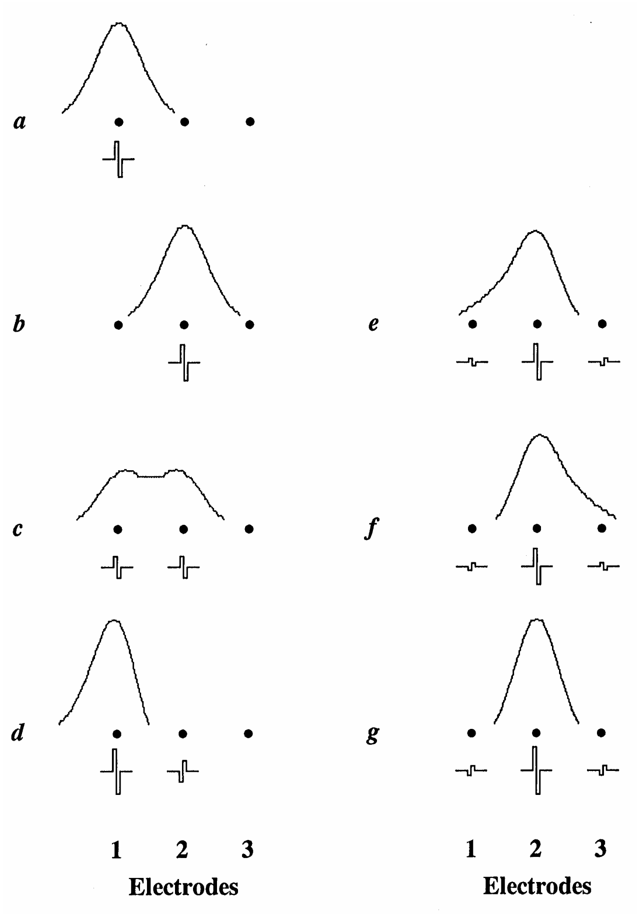

Figure 2. Schematic illustrations of neural responses for various conditions of stimulation. The top curve in each panel is a hypothetical sketch of the number of neural responses, as a function of position along the cochlea, for a given condition of stimulation. The condition of stimulation is indicated by the pulse waveform(s) below each dot, which represent the positions of three adjacent electrodes. Profiles of neural responses for stimulation of a single electrode are presented in *a* and *b*, and profiles for simultaneous stimulation of two electrodes in *c* and *d*. Implant subject SR2, when listening to these different stimuli, balanced for loudness, could rank them according to their distinct pitches. Stimulation of apicalmost electrode 1 alone (*a*) produced a low pitch, whereas stimulation of electrode 2 alone (*b*) produced a higher pitch. Simultaneous stimulation of both electrodes, with identical pulses having approximately half the amplitude of the single-pulse conditions (*c*), produced an intermediate pitch. Pairing stimulation of electrode 1 with a reversed-polarity pulse on electrode 2 (*d*) produced the lowest pitch among the illustrated conditions. Various lower pitches could be produced by manipulating the ratio of the pulse 2 to pulse 1 amplitudes over the range of 0.2 to 0.8. A ratio of 1.0 produced a pitch higher than that elicited by stimulation of electrode 1 alone. Similarly, pitches higher than that elicited by stimulation of electrode 6 alone (the basalmost electrode) could be produced by presenting a reversed-polarity pulse (of lower amplitude) on electrode 5. Additional discriminable pitches between electrodes could be produced by constructing triads of pulses, as illustrated in *e* and *f*. It may even be possible to reduce the width of an unshifted neural excitation field by supplying reversed-polarity pulses on either side of a primary pulse, as illustrated in *g*. Subject SR2 reported that the pitch percept of case *g* was indistinguishable from that of case *b*. The pulse width used for all the listening tests with SR2 was 33 μs/phase. Results from tests with additional subjects using the Ineraid implant are identical to those described above for SR2. In addition, results from those subsequent studies have shown that a relatively-large number of discriminable pitches can be produced between a pair of adjacent electrodes through manipulation in the relative amplitudes of in-phase pulses presented simultaneously to the two electrodes, as in panel *c*, but with different amplitudes for the two pulses. (Illustrations from Wilson et al., 1994.)

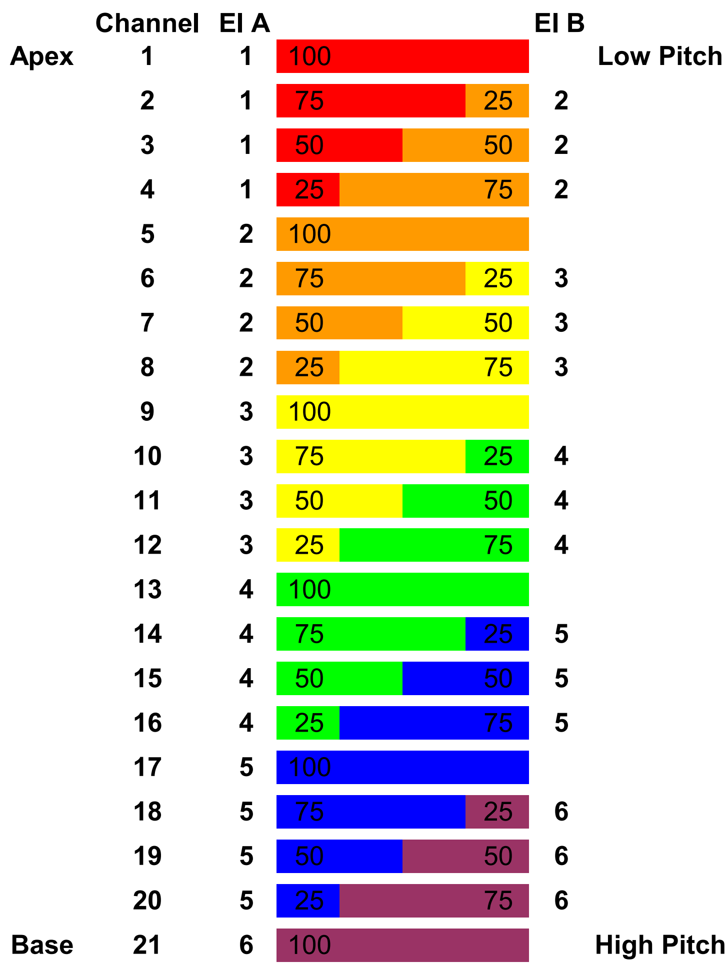

Figure 3. Formal pitch ranking was completed using the two-alternative forced choice paradigm, where two channels are stimulated successively in random order. The subject was asked whether the first or second was higher in pitch. This continues until a statistically significant level is obtained. The ranking order of the channels is indicated from top to bottom and mimics the channel numbering. Statistical significance was obtained with the minimum number of presentations (7 of 7 correct responses) for each of the sequential pairs indicating that these different channels were clearly separable. The virtual channels are represented by the mixed color horizontal bars with the proportions of Electrode A and Electrode B indicating percent contribution of each electrode.

percent steps included 7 comparisons for each pairing; this is the minimum number for statistical significance. Thus, the findings from the tests with 10 percent steps must be regarded as preliminary.)

Although 21 sites were produced by processors using the VCIS approach in the present studies, more sites may be possible, for SR3 and perhaps other subjects as well. This expectation is consistent with the pilot data just mentioned and with the results reported by Litvak et al. (2003). In that latter study, the pitch elicited with simultaneous stimulation of two adjacent electrodes in the Clarion implant (electrodes 2 mm apart, as opposed to the 4 mm spacing in the Ineraid implant) was compared with the pitch elicited by stimulation of the apical electrode in the pair only. For simultaneous stimulation, the proportion of pulse amplitudes was varied in small steps between a relatively high current for the apical electrode to a relatively high current for the basal electrode. Eighteen subjects were tested. The fraction of current needed for the basal electrode to produce a significantly different pitch from stimulation of the apical electrode alone ranged from 1 to 67 percent, depending on the subject. The average across subjects was 19 percent, smaller than the 25 percent steps in current ratios used in the present VCIS implementations, with the Ineraid electrode array (with twice the spacing between adjacent electrodes compared to the spacing in the Clarion array).

The two other processors tested with SR3 included a "standard" VCIS processor as previously described (e.g., Wilson et al., 1994) and a processor that used DRNL filters instead of the linear Butterworth filters used in the standard design. These additional processors, called "VCIS std CIS" and "VCIS cp CIS" respectively, were identical in all other respects except for (1) the exponent used in mapping function for each channel, and (2) an inter-stage gain just prior to the mapping function. In general, the exponent used with cp processors is far higher than the exponent used with standard processors. The higher exponent produces a less-compressive mapping function, that provides an overall compression with cp processors that is similar to the overall compression with standard processors. The different inter-stage gains between the processors are used to provide approximately equal inputs to the mapping function, with quite-different banks of "front end" filters for the two processors.

Methods

Many processors were tested during SR3's most-recent visit. Some of the tests were aimed at further evaluation of parametric choices for cp processors, and other tests were aimed at evaluation of new approaches for representing "fine structure" or "fine frequency" information with cochlear implants. Results from these additional tests with SR3 will be presented in future reports.

The three processors compared in the present report included a max *24*-to-*6* processor using DRNL filters (cp CIS), a 21-site VCIS processor using Butterworth filters (VCIS std) , and a 21-site VCIS processor using DRNL filters (VCIS cp CIS). The cp CIS processor used a max *24*-to-*6* approach, as described above, and was representative of the very best of DRNL *n*-to-*m* processors tested with this subject to date. The cp processors used a less-compressive mapping table than the standard VCIS processor, as also noted above. All processors presented 12 μs/phase pulses in a staggered sequence across the stimulus sites, as is typical for CIS strategies (Wilson et al., 1991). The pulses were presented at each site at the rate of 833 pulses/s. The spectrum spanned by the filterbank was 350-7000 Hz for all processors. The envelope detectors for the processors used a fullwave rectifier followed by a fourth-order, low-pass filter with a Butterworth response and a 200 Hz cutoff frequency. The mapping exponent was -0.0001 for the VCIS std CIS processor (producing a nearly-logarithmic mapping function), and was 0.7 for the cp CIS and VCIS cp CIS processors (producing a less-compressive mapping function). An inter-stage gain was increased in the cp processors to produce an input to the mapping table that was similar to the input for the standard processor. The gain was 4.5 for the VCIS std CIS processor, and was 7.0 or 7.5 for the cp CIS processor, and was 8.0 for the VCIS cp CIS processor. These choices produced speech percepts with each of the processors that were judged by the subject to be comfortably loud, at a level that would be encountered in a normal, everyday conversation.

The tests included identification of 24 consonants in an /a/-consonant-/a/ context and recognition of the City University of New York (CUNY) sentences. The consonants were presented in randomized orders from recordings of multiple exemplars for each consonant by a male talker. Each test included the presentation of each consonant ten times. Each of the sentence tests included the presentation of four separate lists of ten unique sentences from a recorded male talker. Separate sets of lists were used for each of the different tests. The consonants were presented in quiet or in competition with CCITT (speech spectrum) noise at the speech-to-noise ratios (S/Ns) of +5 or 0 dB. The sentences were presented in competition with CCITT noise at the S/N of +5 dB. As noted above, two inter-stage gains were used for the cp CIS processor. A gain of 7.0 was used for the consonant tests, and a gain of 7.5 was used for the sentence tests. The somewhat higher gain for the sentence tests was needed to produce a comfortable loudness for this processor and the recorded CUNY sentences. All tests were conducted with hearing alone and without feedback as to correct or incorrect responses.

Results

Speech reception scores for the three processor conditions are presented in Fig. 4. Analyses of the variance (ANOVAs) indicated significant differences among the scores for the consonant test using the S/N of +5 dB ($F[2,27] = 10.4$, $p < 0.001$), and for the sentence test, also using the S/N of +5 dB ($F[2,9] = 8.8$, $p = 0.008$). The two scores for the consonant test using the S/N of 0 dB are not significantly different ($t[18] = 1.2$, $p = 0.237$). *Post hoc* comparisons using the Holm-Sidak method indicated the following significant differences among the scores for the consonant test at the S/N of +5 dB: VCIS cp CIS is better than cp CIS, and VCIS std CIS also is better than cp CIS. *Post hoc* comparisons using the same method indicated the following significant differences among the scores for the sentence test, also at the S/N of +5 dB: VCIS cp CIS is better than cp CIS, and VCIS cp CIS is better than VCIS std CIS. For each set of comparisons, VCIS cp CIS is better than the other tested processors.

Scores for consonant identification in quiet are greater than 90 percent correct for each of the processors. Ceiling effects may have limited the power of this test to demonstrate possible differences among processors.

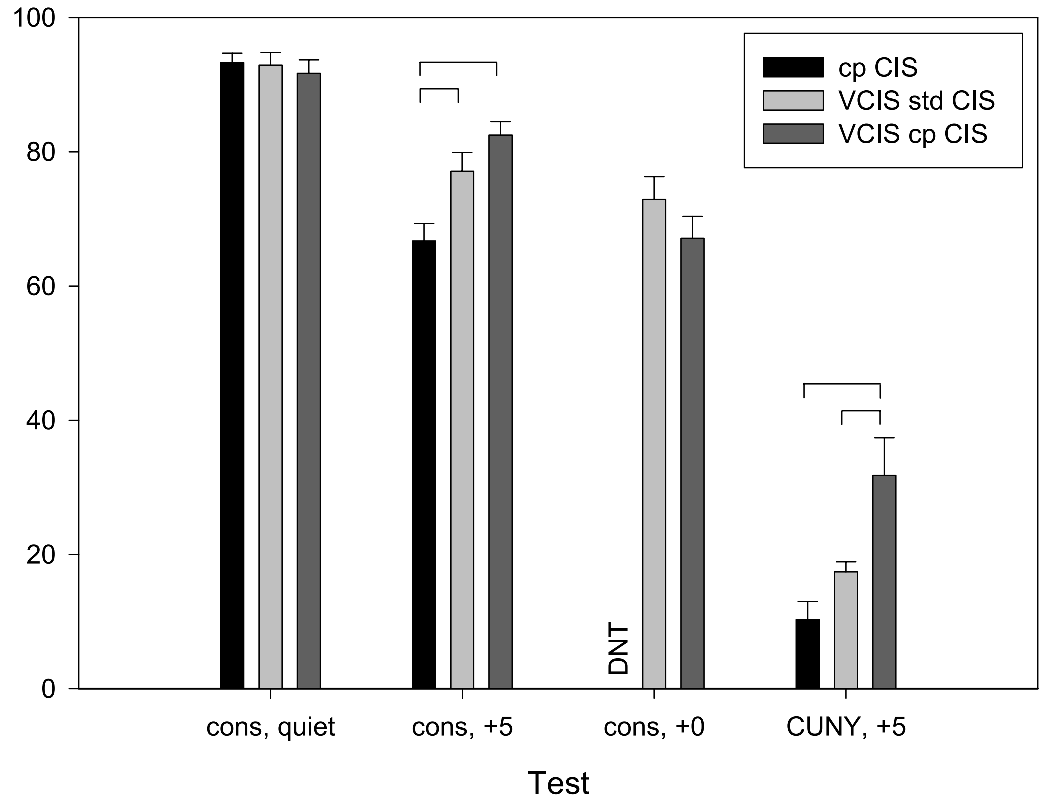

Figure 4. Processor comparisons in tests with Ineraid subject SR3. The processors included a max *24*-to-*6* processor using DRNL filters (cp CIS), a 21-site VCIS processor using Butterworth filters (VCIS std) , and a 21-site VCIS processor using DRNL filters (VCIS cp CIS). The tests included identification of medial consonants in quiet and in noise, and recognition of the City University of New York (CUNY) sentences presented in competition with noise. The speech-to-noise ratios (S/Ns) for the consonant tests included +5 and 0 dB, and S/N for the sentence tests was +5 dB. Analyses of the variance indicated significant differences among the scores for the consonant test using the S/N of +5 dB, and for the sentence test, also using the S/N of +5 dB. The brackets indicate the results of *post hoc* multiple comparisons for those tests, using the Holm-Sidak method. Bars sharing a bracket are significantly different at (at least) the $p < 0.05$ level.

Following completion of the above tests, we wondered how these experimental processors might compare with a standard CIS processor using only six channels of processing and six sites of stimulation. At the end of SR3's visit, a six-channel CIS processor was evaluated using the consonant test at the S/N of +5 dB. The standard CIS processor used the same pulse rate, pulse duration, overall frequency range, and characteristics of the envelope detectors as the experimental processors. The exponent for the mapping function used in the standard CIS processor was the same as that used for

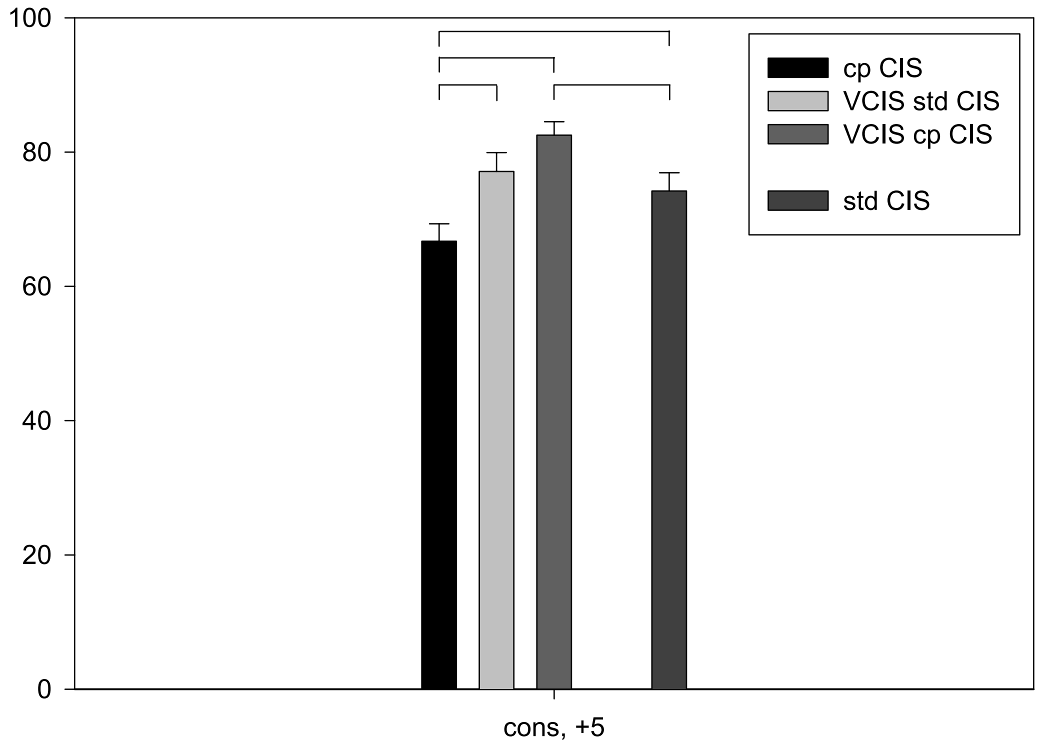

Figure 5. Comparison of a standard CIS processor (std CIS) with the experimental processors included in Fig. 4. The standard CIS processor used six channels of processing and six sites of stimulation. The pulse rate, pulse duration, overall frequency range, and characteristics of the envelope detectors were the same as those for the experimental processors. The exponent for the mapping function was the same as that used for the VCIS std CIS processor. The inter-stage gain for the std CIS processor was 4.0. An analysis of the variance (ANOVA) indicated significant differences among the scores for the only test administered for all four types of processor, consonant identification at the speech-to-noise ratio of +5 dB. The brackets indicate the results of *post hoc* multiple comparisons following the ANOVA, using the Holm-Sidak method. Bars sharing a bracket are significantly different at (at least) the $p < 0.05$ level.

the VCIS std CIS processor. The inter-stage gain for the standard CIS processor was 4.0, the default value for such processors.

The results from the test with the standard CIS processor (std CIS), along with the prior results for the experimental processors, are presented in Fig. 5. An ANOVA indicated significant differences among the scores for the different processors ($F[3,36] = 6.7$, $p = 0.001$). *Post hoc* comparisons using the Holm-Sidak method indicated the following significant differences: VCIS cp CIS is better than cp CIS, and VCIS std CIS is better than cp CIS (as before), and VCIS cp CIS is better than std CIS, and std CIS is better than cp CIS. The VCIS cp CIS is the best processor in these comparisons, and the score obtained with it is significantly better than the scores obtained with the cp CIS or std CIS processors.

Discussion

Among tests and processors, the VCIS cp CIS processor produced the highest scores, and the cp CIS processor produced the lowest scores. For the consonant test at +5 dB, the std CIS processor produced an intermediate score, that is significantly different from the scores for the VCIS cp CIS and cp CIS processors.

These results are consistent with the idea that a relatively high number of (discriminable) stimulus sites may be needed for effective use of DRNL filters in a speech processor for cochlear implants. When a high number of sites is available, possible advantages of DRNL filters may be immediately apparent, as suggested by the present results.

Other strategies, which map a high number of DRNL channel outputs to a small number of stimulus sites (the max *n*-to-*m* or avg *n*-to-*m* approaches), may not be as effective. Combinations or selection of outputs must produce distortions with either approach. However, use of outputs from the DRNL filters, as opposed to the DRNL channels (which include the envelope detectors and nonlinear mapping function), may reduce distortions and thereby improve the performance of *n*-to-*m* approaches. We plan to evaluate this possibility in future studies.

Among the tested processors, SR3 preferred the VCIS cp CIS processor. She asked us to "keep this one," a rare comment for this highly-experienced subject. (This subject also has quite-high levels of performance with her standard CIS processor, which she has used in her daily life for many years.)

As noted above, more than 21 sites may be available using the VCIS approach with SR3 and perhaps other subjects. Future studies should address this possibility. In addition, alternative ways to provide a high number of sites should be investigated. Two such possibilities are to use DRNL processing in conjunction with the Nucleus electrode array, with 22 intracochlear electrodes and up to 22 discriminable pitches for some subjects (e.g., Zwolan et al., 1997), or in conjunction with bilateral cochlear implants, that may provide a higher number of discriminable stimulus sites than either unilateral implant alone (e.g., Lawson et al., 2001).

An even greater number of sites might be produced with a combination of an array with a high number of electrodes and VCIS approach. Of course, this would require simultaneous stimulation of two (or more) electrodes to form the virtual channels. In addition, the number of discriminable virtual channels (or sites) is likely to depend on the spacing of adjacent electrodes. The number may be higher with a wide spacing, e.g., the 4 mm of the Ineraid implant, than with a narrow spacing, e.g., the 0.75 mm spacing of the Nucleus implants. The number of discriminable steps for a given type of electrode array also will without doubt depend on the subject, as has been shown by Poroy and Loizou (2001) for the Ineraid array and by Litvak et al. (2003) for the Clarion "HiFocus" array.

At present, only two implant systems would support the construction of virtual channels in combination with an electrode array with a high number of contacts. These are the Clarion CII device with its HiFocus electrode array and an experimental version of the Nucleus device that includes a percutaneous connector and the "Contour" electrode array. The HiFocus array includes 16 contacts spaced at 1 mm intervals. The Contour array includes 22 contacts spaced at 0.75 mm intervals. The Contour array also has a curved shape, designed to bring its contacts in close proximity to the inner wall of the scala tympani (e.g., Balkany et al., 2002). Such apposition to the inner wall may increase the spatial specificity of stimulation by single electrodes compared with other placements of the electrode array. Both devices support simultaneous stimulation of multiple electrodes.

An important question for future research is how to maximize the number of discriminable sites with cochlear implants*. This might be done with a particular type of electrode array, e.g., one with a high number of contacts in close proximity to the inner wall of the scala tympani, or it might be done using the VCIS approach in conjunction with a particular type of electrode array or with any of a variety of arrays. Alternatively, a dense array of electrodes implanted directly within the auditory nerve may support an especially high number of discriminable sites. (Such intramodiolar implants are under development, see progress reports for NIH project NO1-DC-1-2108, available at http://scientificprograms.nidcd.nih.gov/npp/index.html, and also Badi et al., 2003, and Hillman et al., 2003.)

The present results suggest that a high number of sites, in combination with DRNL processing, may be an especially effective way to represent speech information with cochlear implants. Clearly, studies with additional subjects are needed to evaluate the generality of the present findings with only one subject. In addition, use of different types of electrode arrays, perhaps in combination with virtual channels, may be better than the present use of the Ineraid array in combination with three virtual channels between each set of adjacent electrodes. Our immediate plans include further studies with Ineraid subjects to (1) evaluate the generality of the findings with SR3, using the same processor (and virtual channel) conditions; (2) determine whether more than three discriminable positions may be available between adjacent electrodes of the Ineraid implant; (3) determine the range of variation across subjects (and choices of adjacent electrodes within subjects) in the maximum number of discriminable positions that can be produced using virtual channels; and (4) evaluate *n*-to-*m* approaches that combine or select DRNL filter outputs as opposed to DRNL channel outputs. In these studies, we also plan to investigate in greater detail the parameter space within and across DRNL filters.

---

* A high number of sites may not translate to a high number of useful channels. Results from prior studies have indicated that increases in the number of channels (and electrodes) beyond 4-8 do not produce increases in speech reception scores for users of the standard Nucleus electrode array and the CIS or SPEAK processing strategies (Fishman et al., 1997; Friesen et al., 2001; Lawson et al., 1996; Wilson, 1997). However, results might be different for different types of electrodes, e.g., electrodes in the nerve or in a peri-modiolar position, or for different types of processing strategies, e.g., the VCIS cp CIS strategy of the present report.

In addition to these studies with Ineraid subjects, we also will begin studies in the next quarter with four subjects who now have been implanted (at Duke University Medical Center) with the experimental version of the Nucleus device. Those latter studies will allow evaluation of DRNL processing in conjunction with a high number of stimulus sites, using the 22 electrodes of the Contour array, or using those electrodes in combination with the VCIS approach.

We regard the present findings as encouraging but preliminary. We plan to follow-up with further studies.

References

Badi AN, Kertesz TR, Gurgel RK, Shelton C, Normann RA (2003) Development of a novel eighth-nerve intraneural auditory neuroprosthesis. Laryngoscope 113: 833-842.

Balkany TJ, Eshraghi AA, Yang N (2002) Modiolar proximity of three perimodiolar cochlear implant electrodes. Acta Otolaryngol 122: 363-369.

Fishman KE, Shannon RV, Slattery WH (1997) Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor. J Speech Lang Hear Res 40: 1201-1215.

Friesen LM, Shannon RV, Baskent D, Wang X (2001) Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants. J Acoust Soc Am 110: 1150-1163.

Hillman T, Badi AN, Normann RA, Kertesz T, Shelton C (2003) Cochlear nerve stimulation with a 3-dimensional penetrating electrode array. Otol Neurotol 24: 764-768.

Lawson DT, Wilson BS, Zerbi M, Finley CC (1996) Speech processors for auditory prostheses: 22 electrode percutaneous study – Results for the first five subjects. Third Quarterly Progress Report, NIH project NO1-DC-5-2103, Neural Prosthesis Program, National Institutes of Health, Bethesda, MD. (This report is available online at http://scientificprograms.nidcd.nih.gov/npp/index.html.)

Lawson DT, Wolford RD, Brill SM, Schatzer R, Wilson BS (2001) Speech processors for auditory prostheses: Further studies regarding benefits of bilateral cochlear implants. Twelfth Quarterly Progress Report, NIH project NO1-DC-8-2105, Neural Prosthesis Program, National Institutes of Health, Bethesda, MD. (This report is available online at http://scientificprograms.nidcd.nih.gov/npp/index.html.)

Litvak LM, Overstreet E, Mishra L (2003) Steering current through simultaneous activation of intracochlear electrodes in the Clarion CII cochlear implant: Frequency resolution. Poster presented at the 2003 Conference on Implantable Auditory Prostheses, Pacific Grove, CA, August 17-22, 2003.

Lopez-Poveda EA, Meddis R (2001) A human nonlinear cochlear filterbank. J Acoust Soc Am 110: 3107-3118.

Meddis R, O'Mard LP, Lopez-Poveda EA (2001) A computational algorithm for computing nonlinear auditory frequency selectivity. J Acoust Soc Am 109: 2852-2861.

Poroy O, Loizou PC (2001) Pitch perception using virtual channels. Poster presented at the 2001 Conference on Implantable Auditory Prostheses, Pacific Grove, CA, August 17-22, 2003.

Schatzer R, Wilson BS, Wolford RD, Lawson DT (2003) Speech processors for auditory prostheses: Signal processing strategy for a closer mimicking of normal auditory functions. Sixth Quarterly Progress Report, NIH project NO1-DC-2-1002, Neural Prosthesis Program, National Institutes of Health, Bethesda, MD. (This report is available online at http://scientificprograms.nidcd.nih.gov/npp/index.html.)

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

Wilson BS (1997) The future of cochlear implants. Brit J Audiol 31: 205-225.

Wilson BS, Finley CC, Lawson DT, Wolford RD, Eddington DK, Rabinowitz WM (1991) Better speech recognition with cochlear implants. Nature 352: 236-238.

Wilson BS, Lawson DT, Müller JM, Tyler RS, Kiefer J, et al. (2003) Cochlear implants: Some likely next steps. Annu Rev Biomed Eng 5: 207-249.

Wilson BS, Lawson DT, Zerbi M, Finley CC (1992) Speech processors for auditory prostheses: Virtual channel interleaved sampling (VCIS) processors -- Initial studies with subject SR2. First Quarterly Progress Report, NIH project NO1-DC-2-2401, Neural Prosthesis Program, National Institutes of Health, Bethesda, MD. (This report is available online at http://scientificprograms.nidcd.nih.gov/npp/index.html.)

Wilson BS, Lawson DT, Zerbi M, Finley CC (1994) Recent developments with the CIS strategies. In IJ Hochmair-Desoyer and ES Hochmair (Eds.), *Advances in Cochlear Implants*, Manz, Vienna, pp. 103-112.

Wilson BS, Lawson DT, Zerbi M, Finley CC (1995) Speech processors for auditory prostheses: "Virtual Channel" and "Sharpened Field" CIS processors (among other topics). Final Report, NIH project NO1-DC-2-2401, Neural Prosthesis Program, National Institutes of Health, Bethesda, MD. (This report is available online at http://scientificprograms.nidcd.nih.gov/npp/index.html.)

Wilson BS, Zerbi M, Lawson DT (1993) Speech processors for auditory prostheses: Identification of virtual channels on the basis of pitch. Third Quarterly Progress Report, NIH project NO1-DC-2-2401, Neural Prosthesis Program, National Institutes of Health, Bethesda, MD. (This report is available online at http://scientificprograms.nidcd.nih.gov/npp/index.html.)

Zwolan TA, Collins LM, Wakefield GH (1997) Electrode discrimination and speech recognition in postlingually deafened adult cochlear implant subjects. J Acoust Soc Am 102: 3673-3685.

III. Plans for the next quarter

Among the activities planned for the next quarter are:

- A two-week visit by new subject ME-23, using a Med-El cochlear implant with substantial residual acoustic hearing, January 5-16.
- A visit by investigator Artur Lorens from Warsaw, Poland, to collaborate in studies with this subject.
- Studies with new local subject ME-24, who has bilateral Med-El implants.
- Initial studies with second Nucleus Contour Electrode percutaneous subject NP-8, beginning February 2-3.
- An invited presentation by Blake Wilson in Valencia, Spain, February 22.
- Studies with NP-8 the week of March 1-5, to evaluate potential of new DRNL filter processing strategies in conjunction with percutaneous access to many electrodes.
- A visit by consultant Enrique Lopez-Poveda to collaborate during the visit by this subject.
- An invited presentation by Blake Wilson during visit to UC-Irvine, March 18.
- Additional studies with local subjects.
- Completion of an invited chapter for "Cochlear Implants: A Practical Guide" (Wilson BS: Processing strategies. In *Cochlear Implants: A Practical Guide, Updated Edition*, edited by H Cooper and L Craddock, Whurr Publishing, London).

IV. Acknowledgments

We thank volunteer research subjects ME14, ME16, ME19, ME22, NP6, and SR3, who participated in studies conducted during this quarter.

Appendix 1: Summary of reporting activity for this quarter

Reporting activity for this quarter, covering the period of October 1 through December 31, 2003, included the following invited presentations:

Wilson BS, Guest of Honor Address, Evaluation of combined EAS in studies at the Research Triangle Institute, *Hearing Preservation Workshop II*, Frankfurt, Germany October 17-18, 2003.

Lawson DT and Schatzer R: Speech Processors for Auditory Prostheses. *34th Neural Prosthesis Workshop,* NIH, Bethesda, MD, October 21-23, 2003.

In addition to these presentations, Blake Wilson served as the chair for one of the sessions at the Hearing Preservation Workshop:

Wilson BS (Chair): Session on "Clinical Issues," *Hearing Preservation Workshop II*, Frankfurt, Germany, October 17-18, 2003.