First Quarterly Progress Report

January 1, 2002, through March 31, 2002

Speech Processors for Auditory Prostheses

NIH Contract N01-DC-2-1001

submitted by

Donald K. Eddington
Joseph Tierney
Victor Noel
Margaret Whearty

Massachusetts Institute of Technology
Research Laboratory of Electronics
Cambridge, MA

1.0 Introduction

Work performed with the support of this contract is directed at the design, development, and evaluation of sound-processing strategies for auditory prostheses implanted in deaf humans. The investigators, engineers, audiologists and students conducting this work are from four collaborating institutions: the Massachusetts Institute of Technology (MIT), the Massachusetts Eye and Ear Infirmary (MEEI), Boston University (BU) and the University of North Carolina at Chapel Hill(UNC-CH). Major research efforts are proceeding in four areas: (1) developing and maintaining a laboratory-based, software-controlled, real-time stimulation facility for making psychophysical measurements, recording field and evoked potentials and implementing/testing a wide range of monolateral and bilateral sound-processing strategies, (2) refining the sound processing algorithms used in current commercial and laboratory processors, (3) exploring new sound-processing strategies for implanted subjects, and (4) understanding factors contributing to the wide range of performance seen in the population of implantees through psychophysical, evoked-response and fMRI measures.

A good deal of this first quarter's effort was directed at preparing for future experiments. For instance, the development of the software/hardware to control simultaneous stimulation of two Clarion CII/HiFocus implant systems was the focus of most of the MEEI's engineering work. Experiments with our first, bilaterally implanted subject will begin in the next quarter using the facilities developed in Q1. Similarly, work conducted at the UNC-CH in collaboration with Advanced Bionics Corporation hs resulted in a package of stimulating/recording tools for the Clarion CII/HiFocus implant system that will be used to begin field and evoked-response measures in Q2. These tools have been delivered to the MEEI where field-potential and intracochochlear evoked-potential (IEP) measures have been successfully recorded in our first subject. Details of these stimulation and recording tools will appear in subsequent Quarterly Progress Reports (QPRs) as we begin reporting the results of their use. In this QPR, we concentrate on psychophysical measures of interaction using waveforms designed to reduce the influence that stimulation at one electrode has on a neighbor.

2.0 Background

Figure 1 is an example of an early sound-processing strategy used for cochlear implants (Eddington 1980). After an automatic gain control (AGC), the microphone signal is presented to a set of band-pass filters that separate the sound spectrum into four processing channels. The current sources translate the voltage waveforms at the filters' outputs to the current waveforms delivered to the implanted electrodes. Output channels are connected to electrodes such that the higher the center frequency of a channel's band-pass filter, the more basal its electrode's position.

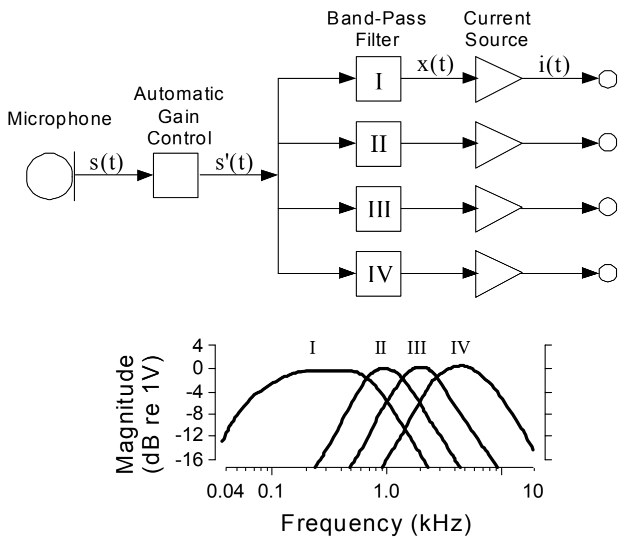

Figure 1. Top: block diagram of an early, four-channel sound processing system. Bottom: magnitude of the band-pass filters' transfer functions.

The dynamic range associated with electric hearing ranges from 3 to 24 dB (Eddington, Dobelle et al. 1978). This means that the 120-dB dynamic range of acoustic hearing must be compressed by the AGC. This system's name, Compressed Analog (CA), stems from the analog nature of the stimulus waveforms and the front-end compression (Wilson, Finley et al. 1991).

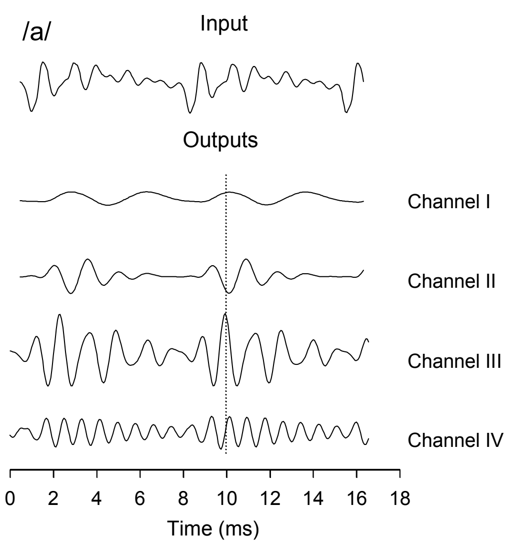

Figure 2. Stimulus waveforms produced by a four-channel CA processor in response to the vowel /a/. The top waveform is the input signal and the four bottom waveforms are the output signals of channels I through IV (see Figure 1).

One problem with the CA strategy is illustrated in Figure 2 where the output waveforms in response to the vowel /a/ are plotted for an Ineraid sound-processing system. Note that the stimulus produced by channel III is relatively strong, indicating significant energy in the input signal within the bandwidth of that channel. The vertical line of this figure marks a time when the output of channel III reaches a peak and channel II is delivering a negative signal. Because the distance between the electrodes of these neighboring channels is approximately 4mm, their potential distributions will overlap and the responses of a significant number of nerve fibers will be influenced by the stimuli of both channels. At this time, the stimuli from these two channels are out of phase and will tend to cancel. This kind of interaction between the stimuli of two or more electrodes represents a distortion that can adversely affect speech reception.

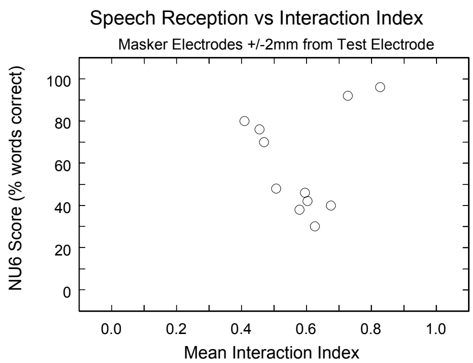

Figure 3. Scatter plot of single-syllable word recognition as a function of the mean Interaction Index for simultaneous stimulation11 CI/HiFocus subjects.

Measures of interaction suggest a negative correlation between a patient's ability to receive speech information and the degree to which a subthreshold masking stimulus delivered to the masker electrode influences the threshold measured on a neighboring test electrode. We defined an interaction index (II) that varies between 0 (no interaction) to 1 (threshold shifted the same amount as when the masker and test stimuli are both delivered to the test electrode) (Eddington and Whearty 2001). Figure 3 shows data that suggest the ability to recognize single-syllable words is negatively correlated with the II in 9 of the 11 subjects.

One approach that reduces interaction is to use a processing strategy that temporally interleaves stimuli across electrodes (Eddington, Dobelle et al. 1978; Wilson, Finley et al. 1991). Two channels of such a processing strategy are shown in Figure 4. Like the CA processor of Figure 2, this processor uses a set of band-pass filters to separate the spectrum into a number of channels. Each channel then extracts the filtered signal's envelope and uses it to amplitude modulate a biphasic pulse train. After compression by a level-mapping function, this modulated pulse train is delivered as a current waveform to the electrode. The pulsatile nature of the stimulus makes it possible to adjust the relative timing of the pulse trains across channels so that only one electrode receives non-zero stimulation current at any one time. This style of signal processing is called a Continuous Interleaved Sampling (CIS) processing strategy (Wilson, Finley et al. 1991).

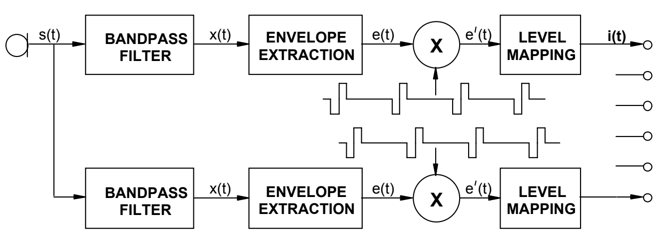

Figure 4. Block diagram of a processing strategy that interleaves stimuli across stimulating electrodes.

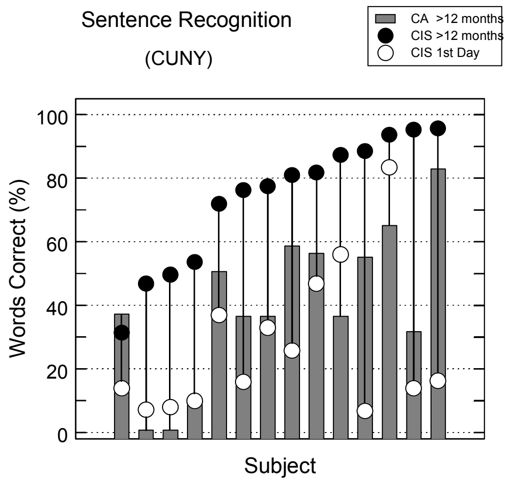

Figure 5. Percentage of words identified correctly when lists of the CUNY sentences are presented without speech reading to 14 profoundly impaired users of the Ineraid cochlear implant system. Each subject was tested at three times: (1) after 12 months experience using a CA style sound processor (bars), (2) the same day they switched from the CA processsor to a CIS processor (open cirlcles), and (3) after 12 months experience with the CIS processor (filled circles).

Figure 5 shows the effect on speech reception of switching from a CA to a CIS strategy in 14 local subjects . Different lists of the recorded CUNY sentences (Boothroyd, Hanin et al. 1985) were used (without speechreading) to evaluate performance of the subjects at the three times described in the caption. These test materials are relatively easy because the internal predictability of each sentence (e.g., “Take your baseball glove to the game.”) enables one to piece together the unrecognized segments from the scattered segments that are recognized.

The bars of Figure 5 represent the word scores of the 14 subjects tested using their CA strategy. At the time of the test, they had worn that system for at least 12 months. The scores for this case range from 0 to 82%. The open circles represent the scores measured using the CIS system on the day subjects switched to this new processing strategy. Note that in some cases performance increased immediately but in others it decreased substantially. After using the CIS strategy for more than 12 months, performance was measured again (filled circles).

3.0 Triphasic Stimulus Waveforms

It is clear that the CIS system resulted in better speech reception for most of these subjects. Because this improvement is likely due (at least in part) to a reduction in the interaction between stimuli delivered by separate electrodes, other techniques for minimizing interaction may also prove beneficial. One method we are currently investigating is optimizing the stimulus waveform. This is motivated by data like those shown in Figure 6 where the model (Frijns 1995) nerve-fiber responses to the two stimuli diagrammed in the

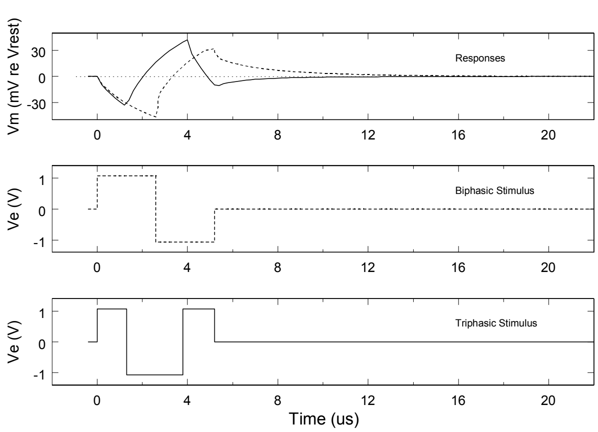

Figure 6. The top panel plots the membrane voltage (Vm) at a single node of a single-fiber model in response to the two electric stimuli shown in the bottom two panels. Ve represents the voltage at the same node.

bottom panels are plotted in the top panel. Note that in the case of the triphasic stimulus, the membrane potential (Vm) is substantially closer to the resting value at the end of the stimulus waveform than in the case of the biphasic waveform. This means that the response to a second stimulus directly following the first is more likely to be influenced by the biphasic than the triphasic waveform.

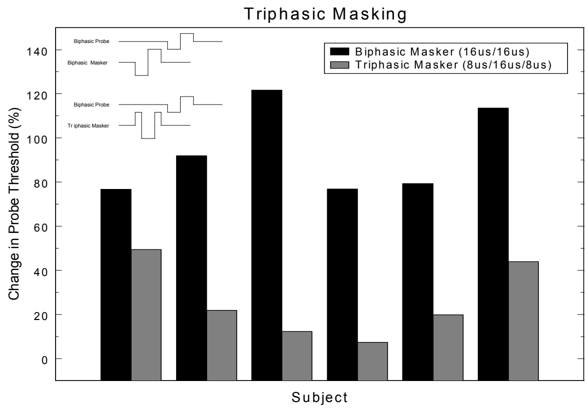

Figure 7. The time waveforms in the upper left show the relationship of the biphasic and triphasic maskers to the biphasic probe waveform. The bars represent the percentage change in the probe threshold when measured in the presence of each masker for six subjects.

The results in Figure 7 show the degree to which the behavioral threshold of a biphasic pulse (16 μsec/phase) delivered to one electrode (the probe stimulus) is influenced by a superthreshold, masker stimulus delivered to an adjacent electrode (4 mm electrode separation) for triphasic and biphasic maskers. Note the substantial reduction in the masker's effect for triphasic vs. biphasic masker waveforms.

4.0 Future Work

Next Quarter we plan to continue work directed at triphasic stimulation waveforms. We are beginning to acutely test speech reception in subjects using CIS sound-processing strategies where the band envelopes modulate trains of triphasic pulses. The speech reception of these CIS strategies will be compared with the classic CIS strategy modulating biphasic pulse trains.

The first of our five monolaterally implanted subjects scheduled to undergo implantation of their second ear will be ready to begin bilateral testing next quarter. Initial experiments with this subject will be directed at testing the bilateral stimulation system we have been working to develop this quarter. We expect to begin measuring this subject's head-related transfer functions for use in future experiments and conducting initial measures of ITD JNDs. Three additional subjects are scheduled to receive their second implants during Q2.

The software developed and tested during Q1 for field and evoked-potential recording from intracochlear electrodes of the Clarion CII/HiFocus implant system will be used to make measures in an initial group of monolaterally-implanted Clarion subjects. The objectives of collecting these initial data are to better characterize system measurement noise, identify software refinements to improve speed and quality of data collection, and to survey the pool of prospective subjects with regard to the magnitude and quality of their IEP measures.

Hardware and software development will continue toward the goal of objectively verifying proper operation of the implant before embarking on extensive data collection. In addition, modification of the IEP measurement tools to enable measurement of channel interaction will be initiated as a compliment to the previously described psychophysical measures of channel interactions.

Boothroyd, A., L. Hanin, et al. (1985). A Sentence Test of Speech Perception: Reliability, Set Equivalence, and Short Term Learning. New York, NY, City University of New York.

Eddington, D. K. (1980). "Speech discrimination in deaf subjects with cochlear implants." J Acoust Soc Am **68**(3): 885-91.

Eddington, D. K., W. H. Dobelle, et al. (1978). "Auditory prosthesis research with multiple channel intracochlear stimulation in man." Ann Otol Rhinol Laryngol Suppl **87**(6 Pt 2): 1-39.

Eddington, D. K. and M. Whearty (2001). Electrode interaction and speech reception using lateral-wall and medial-wall electrode systems. 2001 Conference on Implantable Auditory Prostheses, Pacific Grove, CA.

Frijns, J. H. M. (1995). Cochlear implants: a modelling approach. Faculteit der Godgeleerdheid. Eindhoven, Rijksuniversiteit te Leiden**:** 184.

Wilson, B. S., C. C. Finley, et al. (1991). "Better speech recognition with cochlear implants." Nature **352**(6332): 236-8.