SOUND PROCESSING FOR COCHLEAR IMPLANTS

Donald K. Eddington.[1,2,3] *William M. Rabinowitz,* [4] *and Lorraine A. Delhorne* [2]

[1]Cochlear Implant Research Laboratory, Massachusetts Eye and Ear Infirmary
[2]Research Laboratory of Electronics, Massachusetts Institute of Technology
[3]Department of Otology and Laryngology, Harvard Medical School
[4]Bose Corporation

ABSTRACT

Cochlear implants are devices designed to provide a measure of hearing to the deaf. Most deaf individuals have lost the ability to translate sound into the patterns of electric activity normally present on the 30,000 fibers of the auditory nerve. Because these patterns of activity are the inputs to the brain that result in sound sensation, cochlear implants deliver electric stimuli to these fibers in an attempt to artificially elicit patterns of spike activity that mimic the patterns present in a normal-hearing ear.

We introduce cochlear implants by describing the signal processing used by current devices. Measurements of patient performance in quiet and in noise are used to demonstrate the limitations of today's devices and to introduce the avenues of current research that show promise for improving the performance of these devices.

1. INTRODUCTION

This overview is organized into three sections. First we introduce the rationale for these devices by reviewing the normal process by which acoustic signals are converted to neural activity, examining the disruptions that lead to hearing impairment, and showing how a cochlear implant is designed to overcome impairment.

Next we present an example of how speech reception was improved for a group of patients by altering their sound processing strategy. This example introduces a number of signal-processing issues encountered with cochlear implants and presents data illustrating the range of performance associated with these devices.

Finally, we examine the results of speech-reception tests conducted with a normal-hearing subject listening to an acoustic simulation of a sound processing strategy used by current implantees. These data suggest several factors limiting the performance of today's sound processing strategies.

Supported by NIH grant DC003617, NIH contract N01-DC-6-2100 and the W.M.Keck Foundation

2. RATIONALE

The top panel of Figure 1 shows a schematic diagram of the normal peripheral auditory system. The ear canal and ossicles (small bones) of the middle ear transmit acoustic signals to the cochlea where they produce a travelling wave moving from base to apex along the basilar membrane. Displacement of a segment of the basilar membrane increases the likelihood that the hair cells coupled to that segment will cause their nerve fibers to elicit spikes. The structural properties of the basilar membrane result in a maximal displacement for high

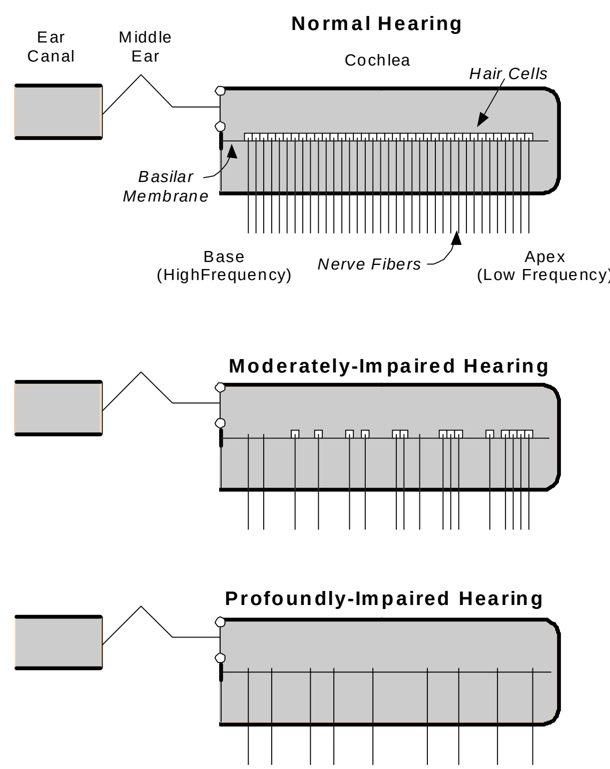

Figure 1. Schematic diagrams of the peripheral auditory system for cases of normal, moderately-impaired and profoundly-impaired hearing.

frequencies at the cochlea's base and for low frequencies at the apex. As a result, the spectral content of an acoustic stimulus is represented by an array of nerve-fiber responses where the highest-frequency components are coded by fibers innervating hair cells at the base and the lowest frequency components at the apex.

The middle panel of Figure 1 represents the case of moderate hearing impairment where some hair cells and nerve fibers have been destroyed. Such an impairment can result from a number of causes like bacterial or viral infection, genetic programming and acoustic trauma. When the number and distribution of undamaged hair cells and nerve fibers support sufficient residual hearing, a hearing aid that amplifies the acoustic signal can provide a good deal of benefit.

Unfortunately, this is not the case for the profoundly impaired. As depicted in the bottom panel of Figure 1, few hair cells are available to excite the remaining nerve fibers and amplification is ineffective.

Cochlear implants are devices designed to use electric stimulation of the remaining auditory-nerve fibers to restore a measure of hearing to the profoundly impaired. The basic structure of the device is diagrammed in Figure 2. An array of electrodes (unfilled circles) are surgically implanted in the cochlea and connected to a sound processor. The sound processor (typically DSP-based) is programmed to translate the output of the microphone into electric delivered to one of the implanted electrodes. The number of processing and stimulation channels range from 4 to 24.

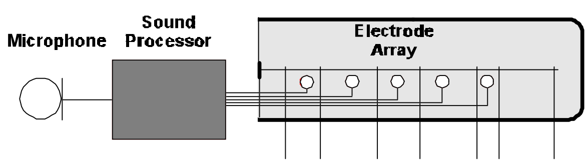

Figure 2. Schematic diagram of a cochlear implant system.

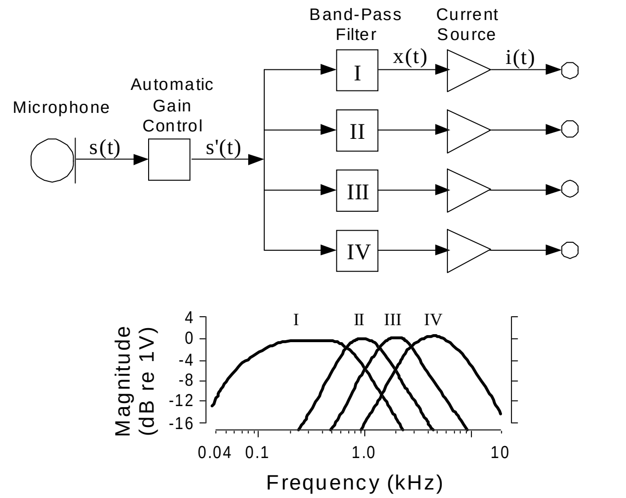

Figure 3. Top: block diagram of an early, four-channel sound processing system. Bottom: magnitude of the band-pass filters' transfer functions.

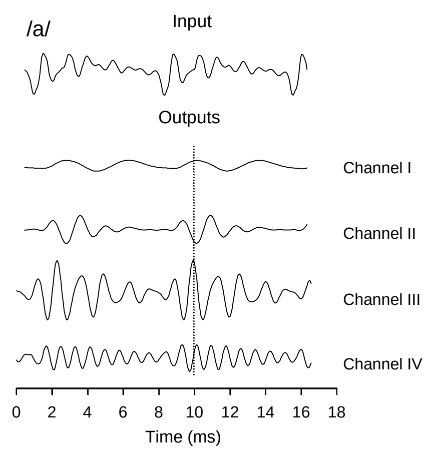

Figure 4. Stimulus waveforms produced by a four-channel CA processor in response to the vowel /a/. The top waveform is the input signal and the four bottom waveforms are the output signals of channels I through IV (see Figure 3).

3. TWO PROCESSING STRATEGIES

The processing strategy shown in Figure 3 is an example of an early strategy used for cochlear implants [1]. 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.

The dynamic range associated with electric hearing ranges from 3 to 24 dB [2]. 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.

One problem with the CA strategy is illustrated in Figure 4 where the output waveforms in response to the vowel /a/ are plotted. 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 less than 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

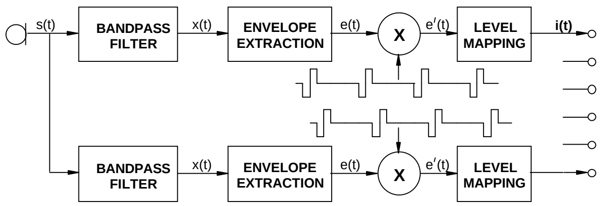

Figure 5. Block diagram of a processing strategy that interleaves stimuli across stimulating electrodes.

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.

One approach that can reduce interaction is to use a processing strategy that temporally interleaves stimuli across electrodes [2, 3]. Two channels of such a processing strategy are shown in Figure 5. Like the CA processor of Figure 3, 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.

Figure 6 shows the effect on speech reception in 14 subjects of switching from a CA to a CIS strategy. Different lists of the recorded CUNY sentences [4] were used (without speechreading) to evaluate performance of the subjects at the three times described in Figure 7's 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 6 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).

It is clear that the CIS system resulted in better speech reception for most of these subjects. Experience tells us that subjects scoring better than ~85% on this task will be able to converse with sufficient fluency to carry on

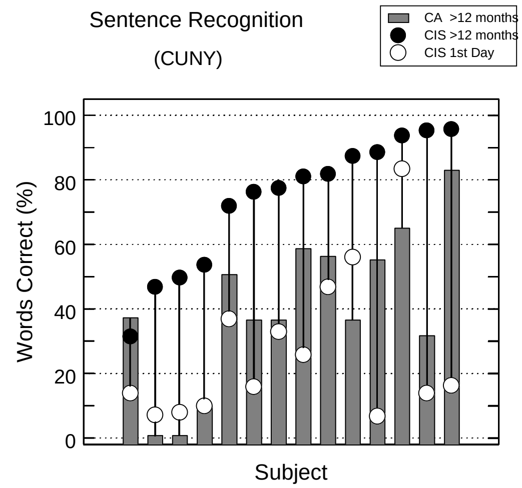

Figure 6. Percentage of words identified correctly when lists of the CUNY sentences are presented without speech reading to 14 profoundly impaired users of the Ineraid cochlear implant system. Each subject was tested at three times: (1) after 12 months experience using a CA style sound processor (bars), (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).

conversations without speechreading (e.g., conduct significant business over the telephone). For lower performing subjects, fluent conversation requires speechreading together with the sound information conveyed by the implant.

Notice also the large range of performance represented by these subjects. Unfortunately, it is impossible to predict before surgery where in this range of performance a particular patient will land.

The decrease in performance measured for many subjects on the day the sound processing strategy was switched illustrates a challenge faced by investigators focused on improving sound processing strategies. Namely, scores from acute testing cannot be used as a reliable metric of a strategy's potential.

4. ACOUSTIC SIMULATION OF A CIS PROCESSING STRATEGY

In an attempt to gain insight into the information implantees derive from CIS sound processing strategies, we developed a signal processing system designed to deliver information to acoustic listeners that is similar to the information received by implantees [5]. As shown in Figure 7, the input signal is initially processed like a CIS processor. Instead of delivering the envelope-modulated pulse train of each channel to a different electrode (cochlear position), the envelope of each channel modulates a tone that directs that channel's envelope information to the appropriate cochlear place of the

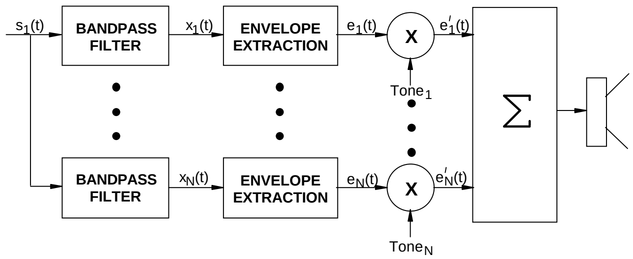

Figure 7. Block diagram of the signal processing system used to acoustically simulate a CIS processing strategy. The band envelope [e(t)] carries the temporal information delivered to an implantee's electrode. The envelope of each channel modulates a tone at the geometric center frequency of the band-pass filter associated with that channel. The modulated tones are summed and played for a normal-hearing listener using headphones or a speaker [6].

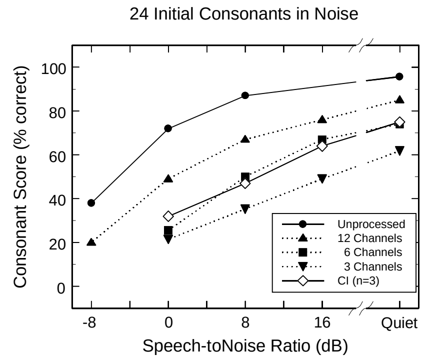

Figure 8. Measures of initial consonant reception as a function of speech-to-noise ratio for five conditions. Filled symbols represent scores for one normal-hearing subject. Scores measured without any processing are shown by filled circles. Upward-pointing triangles, squares and downward-pointing triangles represent scores measured using the CIS simulator shown in Figure 7 with the speech spectrum split into 12, 6 and 3 channels respectively. The open diamonds are mean scores for three of the best performing implant subjects using CIS processors of 6 or 8 channels.

acoustic listener. The sum of these amplitude-modulated tones is the acoustic output of the simulator. Note that the simulation does not include a compressive mapping function because the small dynamic range of electric hearing is not an issue for normal-hearing listeners.

We used the simulation system of Figure 7 to explore the effect of the number of processing channels on a subject's ability to recognize the 24 initial consonants of English when presented in a consonant-vowel-consonant context in both quiet and in noise (for details see [6]). The results of Figure 8 show an orderly decrease in the normal-hearing subject's performance (filled symbols) as the number of channels and the speech-to-noise ratio (SNR) decrease. The mean scores for the three high-performing cochlear implant users also decrease as SNR decreases.

One interesting feature of these data is the close correspondence between the scores of the implant subjects (6 or 8-channel processors) and the normal-hearing subject listening through a 6-channel simulation. If one assumes that the normal-hearing listener extracts virtually all the information available in the signal of the 6-channel simulation, this means that the high-performing implantees are also extracting virtually all of the information relevant to speech that is available from their implant. This suggests that further gains in performance can be obtained only by increasing the number of effective CIS channels/electrodes or by altering the processing strategy to provide additional information within the constraints of the existing electrodes.

5. SUMMARY

The benefit deaf adults receive from cochlear implants varies widely across patients. The top 20% are able to converse quite fluently without the aid of speechreading. Virtually all implantees are able to combine speechreading cues with the information provided by the implant to converse much more fluently than with a hearing aid.

Improvements in the performance provided by today's devices will require new systems that substantially increase the number of effective information channels or increase the information effectively encoded in the existing channels.

The oral presentation will address current efforts to move forward on both of these fronts.

6. REFERENCES

[1] D.K. Eddington, "Speech discrimination in deaf subjects with cochlear implants," *J Acoust Soc Am*, vol. 68, pp. 885-91, 1980.

[2] D.K. Eddington, W.H. Dobelle, D.E. Brackmann, M.G. Mladevosky, and J. L. Parkin, "Auditory prosthesis research with multiple channel intracochlear stimulation in man," *Ann Otol Rhinol Laryngol Suppl*, vol. 87, pp. 1-39, 1978.

[3] B.S. Wilson, C.C. Finley, D.T. Lawson, R D. Wolford, D.K. Eddington, and W.M. Rabinowitz, "Better speech recognition with cochlear implants," *Nature*, vol. 352, pp. 236-8, 1991.

[4] A. Boothroyd, L. Hanin, and T. Hnath, "A Sentence Test of Speech Perception: Reliability, Set Equivalence, and Short Term Learning," City University of New York, New York, NY, Speech and Hearing Science Report RC110, 1985.

[5] RV. Shannon, F-G. Zeng, V. Kamath, J. Wygonski, and M. Ekelid, "Speech Recognition with Primarily Temporal cues," *Science*, vol. 270, pp. 303-304, 1995.

[6] D.K. Eddington, W.R. Rabinowitz, J. Tierney, V. Noel, and M. Whearty, "Speech Processors for Auditory Prostheses, Eighth Quarterly Progress Report," MIT, Research Laboratory of Electronics, Cambridge December 1997.