The Feasibility of a Cochlear Nucleus Auditory Prosthesis Based on Microstimulation

NIH Contract #N01-DC-5-2105

Quarterly Progress Report #11

January 1, 1998 to March 31, 1998

Robert V. Shannon
Jean Moore
House Ear Institute
2100 W. Third St.
Los Angeles, CA 90057

Douglas McCreery
William Agnew
Ted Yuen
Leo Bullara
Huntington Medical Research Institutes
734 Fairmount Ave.
Pasadena, CA 91105

INTRODUCTION

The work statement of the present contract calls for the development of psychophysical protocols for evaluating the performance of auditory brainstem implants (ABI's) penetrating microelectrodes within the ventral cochlear nucleus. This quarterly report describes issues related to the evaluation of ABI patients performance and to development of speech processor strategies.

Prosthetic electrical stimulation of the auditory system has been far more successful than had originally been predicted. With modern implant devices, most deaf people with cochlear implants can understand speech well enough to converse over the telephone. Recently, prosthetic electrical stimulation has been applied to the cochlear nucleus of patients with no remaining auditory nerve (Brackmann et al, 1993). Although the speech recognition performance of these patients is not as good as in patients with multichannel cochlear implants, the auditory brainstem implant (ABI) provides clear benefit to deaf patients (Shannon et al., 1992, 1993; Otto et al, 1998).

Initially, auditory researchers assumed that artificial (electrical) stimulation of the auditory system would not be very useful because the pattern of neural activity would be highly unnatural. Electrical stimulation cannot recreate the complex temporal and spatial patterns of neural activity that exist in a normal ear. However, cochlear implants, while producing an unnatural and artificial pattern of nerve activity, do allow for considerable speech recognition. We now know that the reason for this apparent discrepancy is the powerful pattern recognition capability of the human brain. A sensory prosthesis does not have to reproduce the all of the details of the peripheral activity pattern to be useful. While we understand the details of the peripheral pattern quite well, and while those details may play a role in fine auditory discriminations, they are clearly not all necessary for recognition of speech. Because listeners must be able to understand speech in a broad range of acoustic conditions and by a wide range of speakers and speaking styles, speech pattern perception is largely a higher-order cognitive function not critically dependent on all details of the peripheral pattern of nerve activity. The main task in the development of auditory prostheses is to determine <u>which</u> global cues in the sensory peripheral response patterns are most important for cuing the central pattern recognition process and how to convey <u>those</u> cues most effectively.

Present ABI Technology

Present technology produces electrical stimulation of the human cochlear nucleus through surface electrodes placed in the lateral recess of the IV ventricle (adjacent to the dorsal cochlear nucleus: DCN) at the time of surgery to remove a vestibular schwannoma (Brackmann et al, 1993; Shannon et al., 1993; Otto et al., 1998). The present surface array contains eight electrodes, and it appears that the use of multiple electrodes produces significantly better performance than do the earlier single-channel systems, at least in some patients (Otto et al., 1998).

The function of the DCN is not well understood, but it is thought to have a general alerting function, and possibly some role in sound localization (Masterton, 1992). Unfortunately, the tonotopic axis of the DCN is orthogonal to the axis of the surface electrode array. Thus, even if each electrode were stimulating a distinct group of fibers, we might not expect the electrodes to elicit different pitches. And because of the hypothesized specialized nature of processing in the DCN, several prominent auditory physiologists have speculated that the sensations evoked would be of little value for speech understanding. Cats who have had their DCN lesioned show no clear deficits in auditory function (Masterton, 1992)

In spite of the apparently inappropriate and non-tonotopic electrode location, the present surface electrode system has produced useful prosthetic benefit. ABI patients receive significant improvement in lipreading, and can distinguish sounds on the basis of temporal envelope cues (Shannon et al., 1993). Multi-electrode ABIs produce electrode-specific pitch in most patients, although the pitch map is not consistent across patients (Shannon et al., 1993; Otto et al., 1998). Significant open-set speech recognition has been achieved in about 10% of multichannel ABI patients (Otto et al., 1998) although this level of performance is not as good as that observed with cochlear implants. At the present time it is not clear if multichannel electrical stimulation of the human cochlear nucleus is capable of producing performance comparable to that seen in cochlear implants. The reason for the limited success to date is not clear. Is the poor ABI performance due to the fact that we are stimulating the DCN rather then the VCN? Is it because we are not stimulating the CN in a meaningful tonotopic pattern, or because essential intrinsic processing in the CN is bypassed? Penetrating microelectrode stimulation could overcome some, but not all of these potential limitations.

The first array of penetrating microelectrodes for human use will be composed of 4 or 5 discrete iridium microelectrode of various lengths, to access different point on the tonotopic axis of the posteroventral cochlear nucleus. The array will be part of a hybrid device, which will include the 8-electrode surface array of the extant ABI, the arrays of microelectrodes, and the sound processor-receiver system developed and manufactured by Cochlear Ltd.

If the limiting factor in ABI performance is that the surface electrode is stimulating the wrong subunit of the CN and/or stimulating in a non-tonotopic manner, then a penetrating microelectrode could overcome these problems by stimulating the VCN locally and tonotopically. If the limiting factor is that direct electrical stimulation bypasses or interferes with intrinsic neural processing within the CN, then the same limitation might apply to penetrating electrodes.

Speech Pattern Recognition

Historically, speech recognition research has focused on the role of spectral patterns, e.g. formants and formant transitions, in distinguishing phonemes. Physiological studies have also focused on frequency-based speech cues. Recent speech recognition results in cochlear implant patients (Fishman et al., 1997), and acoustic experiments with reduced spectral cues (Shannon et al., 1995) have shown that high levels of speech recognition are possible with only minimal spectral cues. Only three broad spectral bands (or electrodes) are necessary for 90% correct recognition of words in sentences in quiet.

How much spectral information is necessary to understand speech? Must spectral information be presented to the correct tonotopic region of neurons to be useful? Is the brain flexible enough to utilize the patterns of speech information even if it is coupled to the tonotopic dimension of the auditory nervous system in a nonlinear manner? These questions are as yet unanswered and the answers are critical for the optimal design of auditory prostheses. These questions are even more important in the application of the auditory brainstem implant than in cochlear implants.

In cochlear implants, the electrodes are located in the scala tympani in the basal turn of the cochlea. Electrodes are generally aligned along the tonotopic axis of the nerves. Most researchers assume that successful speech recognition is possible only if the electrodes excite distinct nerve populations that produce distinctive pitches. Consider two hypothetical cases of patients who differ in nerve survival.

Consider the "best" case in which the implant patient has a complete survival of auditory nerves and their peripheral processes (dendrites) and in which the array of electrodes is well positioned relative to the nerve. In this case, each electrode would activate a distinct group of neurons and each electrode would yield a distinct and unique pitch percept. The envelope information from several frequency bands is used to modulate pulses that are applied to the electrodes. In this case each sector of neurons receives a portion of the envelope information from speech. The sectors are properly ordered, in that the envelope from lower-frequency spectral regions are presented to more apical neurons. However, even in this case the neurons are not excited by the envelope information that is appropriate for their "characteristic frequency". The lowest-pitch electrode is located at the "characteristic frequency" location of about 1 kHz. Thus, envelope information from low-frequency region of speech is presented to neurons that would normally respond to considerably higher frequencies. Even so, most cochlear implant patients understand speech quite well.

Now consider the case of a patient with poor nerve survival and/or poor placement of the electrode array. In this case, most of the electrodes may activate the same population of neurons. The patient may perceive little if any pitch difference across the array of electrodes. The pitch elicited by an electrode in any location can be high or low, depending on the "characteristic frequency" of the remaining nerve fibers. In this case, the patient would probably have limited capacity to recognize speech, and the actual nerve fibers stimulated would depend on the location of the surviving nerves, the position of the electrodes, and the current flow pathway. If the current spreads into the modiolus, an electrode carrying high-frequency envelope information could activate low-frequency fibers coming down from the apical end of the cochlea. We assume that a patient in this case would not recognize speech well, but it is not clear which factor (non-homogeneous tonotopic mapping, broad overlap/interaction between electrodes, etc.) would contribute most heavily to the poor performance. It is not yet clear to what extent speech processor adjustments (e.g., re-ordering electrodes in pitch order to compensate for a non tonotopic electrode pitch) can improve speech recognition in such a case.

Now consider electrical stimulation in the cochlear nucleus. The cochlear nucleus contains many anatomically and functionally distinct subunits. Each subunit has its own tonotopic organization, and little is known about the importance of the intrinsic processing in each subunit. In addition, the orientation of the tonotopic axis of each subunit is different from the others. At this time, it is not clear which subunit would be the preferred target region for a penetrating microelectrode, although the posteroventral cochlear nucleus (PVCN) is the leading candidate. An additional concern for penetrating microelectrodes, even if they can activate the tonotopic organization in a selective manner, is that direct electrical stimulation of the cochlear nucleus might bypass some intrinsic neural processing that is essential for speech. Even in the best case, penetrating microelectrodes may not activate evenly spaced tonotopic regions. In an individual patient, for example, most electrodes might stimulate a relatively low-pitch region, while only a few might stimulate higher-pitch regions. Speech processors would probably need to accommodate this non-uniform electrode spacing to optimize speech performance.

Issues in Speech Processor Fitting

Speech processors for auditory prostheses provide a large number of parameters, which must be adjusted as part of the fitting process. Research with cochlear implants has suggested that nonsimultaneous pulses are important for avoiding electrical current field interactions that can occur with simultaneous activation of two or more electrodes. Thus, most present implant speech processors use high-rate nonsimultaneous pulses on multiple electrodes. Most recent research on speech processor parameters have focused on amplitude mapping functions (Boex et al., 1997; Fu and Shannon, 1998b; Shannon et al., 1992), temporal processing (Rosen, 1991; Van Tasell et al., 1987), electrode interaction (Chatterjee and Shannon, 1998; Fu et al., 1996, 1997; Hanekom and Shannon, 1998), the number of electrodes (Dorman et al., 1997; Fishman et al., 1997; Hill et al., 1969), and the tonotopic mapping between frequency information in speech and the electrode location (Dorman et al., 1997; Fu and Shannon, 1998a; Shannon et al., 1998). Each is discussed briefly below.

Amplitude Mapping. Although speech recognition in normal-hearing listeners is relatively robust to amplitude nonlinearities, brainstem implant and cochlear implant listeners have dramatically reduced amplitude ranges and so proper mapping of loudness from acoustic to electrical domains might be more important in this case. Several recent studies have measured speech recognition with custom loudness functions for each electrode (Boex et al., 1997) or under parametric variation in the loudness mapping function (Fu and Shannon, 1998b; Shannon et al., 1992). These studies found that speech recognition was only mildly affected by relatively large variations in the acoustic-to-electrical loudness mapping functions (10-15% decrement in vowel recognition). Thus while the loudness function may contribute to overall quality and naturalness of a speech processor, it appears to be a relatively minor factor in speech recognition performance.

Temporal Resolution. There has been a long-standing discussion about the temporal resolution of hearing-impaired and cochlear implant listeners. The raw psychophysical data indicates that impaired listeners have poorer temporal resolution than unimpaired listeners. However, when the data are plotted in terms of equal loudness, i.e. corrected for the hearing loss, the measures of temporal performance look quite normal. At present, the dominant view is that temporal processing is relatively normal in hearing impaired and implant listeners, once the amplitude is properly mapped to preserve loudness. Indeed, part of the effect observed by loudness manipulations may be due to the fact that extreme amplitude mappings distort the temporal information.

Electrode Interaction. One factor contributing to the reduced speech recognition ability of hearing-impaired listeners is thought to be poor frequency resolution (as evidenced by wider psychophysical tuning curves). Good spectral selectivity is thought to improve the spectral analysis of the speech signal, allowing finer frequency and formant discrimination. In electrical auditory prostheses the spectral (or tonotopic) resolution is determined by the number of electrodes and by the interaction of those electrodes. In a cochlear implant patient with poor nerve survival or in an ABI patient with poor electrode placement, even a large number of electrodes may be so highly interactive that they function as if they were carrying only a single channel of information. Thus, it is important to assess the degree of interaction between electrodes in a cochlear implant and auditory brainstem implant. Over the past 2 years (funded by grant R01-DC-1526) we have developed five methods for quantifying the degree of interaction between electrodes in cochlear implants: pitch estimation (Hanekom and Shannon, 1996; see also Collins et al., 1998; Cohen et al., 1996a, 1996b), electrode discrimination (Hanekom and Shannon, 1996; see also Collins et al., 1998 and Zwolan et al., 1998), forward masking (Chatterjee and Shannon, 1998; see also Shannon, 1983; Lim et al., 1989, and Cohen et al., 1996b), gap detection (Hanekom and Shannon, 1998), and loudness summation (Fu et al., 1996, 1997; see also Shannon, 1985). These methods are all applicable to ABI devices as well. Several of these methods will be tested in the present surface-electrode ABI patients in the next two years to determine the method most suitable to clinical application. Measures of electrode interaction will then be compared between surface electrodes and penetrating electrodes.

Number of Electrodes. It is widely assumed that speech recognition should improve as the number of channels of information is increased, although recent data show that there is a point of diminishing returns between 4 and 8 channels (Dorman et al., 1997a; Fishman et al., 1997; Shannon et al., 1995). In more challenging listening conditions, such as in background noise, more channels are necessary to maintain performance at high levels (Delhorne et al., 1997; Fu et al, 1998). Interestingly, cochlear implant listeners appear to utilize only 4-7 channels of information in speech, even though their devices have as many as 20 electrodes, demonstrating that there is not a one-to-one relation between the number of electrodes and the number of channels of usable information. However, considerable open-set speech recognition is possible with as few as 3-4 channels (Hill et al., 1969; Shannon et al., 1995; Dorman et al., 1997b), so that implant patients are performing well with 4-7 channels, even if this level is sub-optimal relative to the capabilities of the device. Based on these findings, out first intranuclear array will contain 4 or 5 penetrating electrodes; to provide what we estimate to be the best compromise between performance, the risk of tissue injury during electrode insertion, and the flexibility of the electrode cable.

Tonotopic Mapping. In the past few years our lab has obtained new data suggesting that the most important parameters for implant speech processors are: (1) matching the absolute tonotopic location of electrode, and (2) matching the tonotopic extent of electrode, as explained below.

In cochlear implants the electrodes in the scala tympani are located from about 7-25 mm from the round window. This corresponds to tonotopic "best" frequencies of about 500-5000 Hz. However, the speech input to the processor is divided into approximately equal logarithmic frequency bands from 250 Hz to 6000 Hz. This means that the information presented to the most apical electrode may be an octave lower in frequency than the tonotopic location of the electrode. In many cases the electrodes are not inserted fully, resulting in an even larger mismatch between the information presented to an electrode and its tonotopic location. Recent experiments (Fu and Shannon, 1998a; Dolman et al., 1997b) in both normal-hearing listeners and cochlear implant listeners show this to be a highly critical parameter. Indeed, earlier research with frequency compressed or frequency shifted speech (Daniloff et al., 1968; Nagafuchi, 1976; Tiffany and Bennett, 1961) has shown that spectral information in speech can be shifted by about 60% (about 3 mm in tonotopic coordinates) without serious degradation in intelligibility. However, shifts larger than 60% resulted in large decrements, so that by a shift of an octave (4-5 mm) speech recognition was almost completely eliminated. Results by Fu and Shannon show almost the same quantitative pattern in cochlear implants and in normal-hearing listeners simulating cochlear implants. The convergence of data from these diverse studies suggests that the alignment of the spectral information in speech with the proper tonotopic location is a critical factor in speech recognition.

Even if the spectral information is presented to the proper tonotopic location, the homogeneity of this mapping can play a major role in speech recognition. Consider a cochlear implant with 20 electrodes equally spaced along the scala tympani. We might think that if the filters extracting information from speech are matched to the tonotopic location of the electrodes, good speech recognition would result. However, due to the specific pathology leading to deafness of each individual patient, the underlying nerve survival in the spiral ganglion might be uneven along the length of the cochlea. This uneven (or nonhomogeneous) nerve survival would distort the uniformity of the mapping between the spectral representation of speech and the pattern of neural activation. For example consider a case in which no neurons were present in the tonotopic region from 1-2 kHz. The electrodes in that region would still carry the information from speech between 1-2 kHz, but the neurons activated with this information would be at cochlear locations on the edge of this "hole" in nerve survival, i.e., either below 1 kHz or above 2 kHz. This "hole" in the nerve survival would have the effect of warping the spectral representation in the nerve. Experiments in the laboratory with cochlear implant listeners and with normal-hearing listeners simulating cochlear implants have shown that this kind of warping can seriously degrade speech recognition (Fu and Shannon, 1998a).

In an ABI, such matching of the absolute tonotopic location and spacing is much more difficult because we cannot be certain which tonotopic subunit in the cochlear nucleus we are stimulating. Procedures must be developed that will allow us to measure or estimate the absolute tonotopic location of the ABI electrodes and adjust the speech processor accordingly.

Studies to be conducted

Overview A series of measures is proposed to prepare for evaluation of patients with penetrating microelectrode ABI devices, and to determine procedures for optimizing the performance of these devices. Evaluation of basic psychophysical abilities and speech performance with existing surface electrode ABI devices establishes a baseline of performance against which the new device can be compared. Several techniques must be developed for comparison of surface and penetrating electrodes, and for speech processor fitting.

Measures of Performance with ABI Surface and Penetrating Electrodes

Penetrating microelectrodes entail additional risks beyond those of the present surface electrodes. The most likely adverse outcome is damage to an intranuclear blood vessel, which will induce a local microhematoma, and render the device ineffective. Performance of a system based on penetrating electrodes, or on a combination of surface and penetrating electrodes must be superior to that of a system based on surface electrodes alone, in order. to justify the additional risk. Thus, it is important to document the performance levels and capabilities of patients with surface electrode ABI devices, to use as a baseline for assessing improved performance.

Of particular interest is the distinctiveness of the psychophysical percepts from each electrode. We have measured pitch scaling and discriminability of different electrodes in patients with the present surface-electrode ABI system (Otto et al., 1998). Measures that have been developed to assess the overlap in the neural population stimulated by each electrode will be applied in ABI patients. Channel interaction will be measured as a function of electrode configuration and separation.

Speech recognition performance will be assessed in all ABI patients with surface electrodes, and later in patient with the hybrid surface-penetrating system, according to the protocol approved by the FDA. The patient's recognition of consonants, vowels and sentences will be measured in lipreading only, ABI only, and lipreading plus ABI conditions.

Acoustic Simulation of Auditory Prostheses

An acoustic simulation of an auditory prosthesis was recently developed (Shannon et al., 1995) in an attempt to quantify the effects of limited spectral resolution and misplaced spectral representation on speech recognition. The subjects in these experiments had normal hearing, and the signal was delivered acoustically. First, the envelope was extracted from broad spectral bands of the test signal, such as running speech. The envelope was used to modulate band-limited noise, which presumably excited the subjects' cochlear partition corresponding to the bandwidth of the noise. This simulated the loss of spectral detail caused by an electrode stimulating a broad region of nerve fibers in a patient with an auditory implant. The noise bands were then manipulated to simulate different conditions of surviving nerve and electrode placement.

Since the ABI electrodes may be positioned at unequal tonotopic intervals, we will simulate the ABI by 2-8 bands of noise, each modulated by the envelope of a limited frequency band from the speech spectrum. We will simulate non-uniform tonotopic spacing of electrodes by noise bands that are unequally spaced in log frequency (or mm in tonotopic space). Speech recognition performance will be measured for analysis filters that divide the speech spectrum into equal logarithmic parts (like present cochlear implant speech processors), and for analysis filters that match the irregularity of the noise carrier bands. Thus, in the simulation, the analysis band corresponds to a portion of the acoustic spectrum, and the carrier band that it modulates corresponds to a segment of the tonotopic axis of the auditory system. Our hypothesis, based on similar conditions (Fu et al., 1998), is that best performance is obtained when the analysis and carrier bands are matched in terms of tonotopic location and extent. These experiments will quantify the effects of mismatching the analysis bands and carrier bands to simulate the effects of mismatching the spectral division and the tonotopic location of ABI penetrating electrodes. For a brainstem implant, the individual implanted electrodes (which correspond to the individual carrier bands in the simulation) are in fixed locations, so that if the hypothesis is correct we must adjust the location and width of the analysis bands to match the electrodes for optimal performance.

If the penetrating microelectrodes stimulate localized regions of the ventral cochlear nucleus, the resulting pattern of tonotopic activation could be patchy and unevenly spaced. Figure 1 presents a schematic example of the tonotopic representation that might occur with these electrodes.

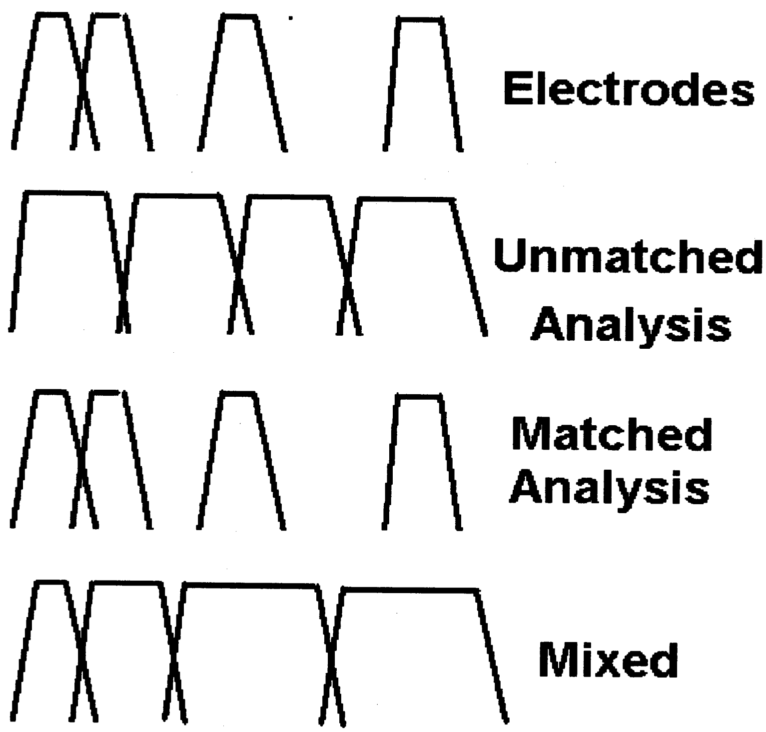

Figure 1

The top row of the figure indicates a possible tonotopic pattern of neural activity that might occur with four penetrating microelectrodes. The extent of activation in tonotopic space (or log frequency) is the same for each electrode and relatively narrow, assuming that penetrating electrodes primarily activate a small population of neurons around the electrode site. However, the spacing of the tonotopic regions may be irregular, due to the uncertainty of access to the tonotopic axis. Our preliminary results suggest that the best speech performance will be realized when the analysis bands are matched to the tonotopic location and bandwidth of the neural populations activated ("Matched Analysis"). At present, the standard clinical speech processor would divide the spectrum into four equal parts (in log frequency) and the speech envelope from each analysis band would be routed to one electrode ("Unmatched Analysis"). A possible intermediate condition is indicated as "Mixed", in which the whole spectrum is represented, but it is divided unevenly to match the relative distribution of the electrodes.

We will simulate these conditions in normal-hearing listeners using narrow bands of noise to simulate the electrodes. The analysis filters will be changed to simulate the "Matched", "Unmatched", and "Mixed" conditions diagramed above to quantitatively assess the consequences of mismatching the analysis filters and electrode locations.

Many patients with the surface electrode ABI report little change in pitch across the electrodes. We suspect that this lack of pitch change will not be a problem with penetrating microelectrodes that access the tonotopic axis in a more localized fashion. However, to prepare for the possibility of a small range of pitch across the penetrating electrodes we will simulate this case to quantitatively assess its impact on speech recognition. We will also measure speech recognition performance with several alternative processor designs to see if performance can be improved by judicious choice of processor parameters.

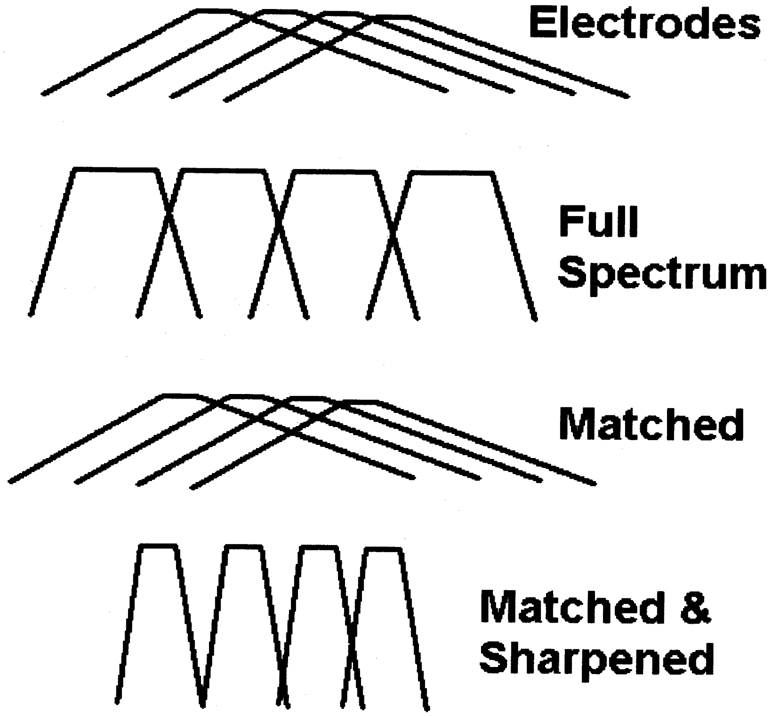

Figure 2

Figure 2 shows three conditions schematically. The top panel sketches the spread of neural activation for four electrodes that are close in tonotopic location and activate broad regions of nerve. Such a representation might produce a small change in pitch across the four electrodes. The second panel illustrates the present division of spectral information used in standard speech processors: the full spectrum is divided into bands that are equally spaced in terms of tonotopic distance (or mm in the cochlea). The bottom panel in Figure 5.3B indicates a processor condition in which the tonotopic extent of the electrodes is matched in the analysis bands of the processor, but the analysis bands are sharper and have less overlap than the electrode excitation patterns. This last condition will test the hypothesis that additional selectivity can be produced by high selectivity on the analysis side, even though the electrodes themselves are highly overlapping.

Low-Rate Speech Processors

Physiological results with penetrating microelectrodes in the cochlear nucleus of cats (McCreery et al., 1997) have demonstrated a stimulation-induced depression of neural excitability (SIDNE). This depression of excitability becomes quite severe at pulse rates of 500 Hz or greater, and the depression can last for many days following termination of stimulation. Although SIDNE may occur in the absense of histologically-detectable tissue injury, severe SIDNE may impair the functionality of the device. Present speech processors utilize nonsimultaneous (interleaved) stimulation on multiple electrodes. Rates of 250 pps to 1000 pps, and even higher rates are proposed by Blake Wilson for cochlear implants. If SIDNE indicates a stimulation regime that should be avoided, the high pulse rate speech processors would not be suitable for microelectrode stimulation of the cochlear nucleus. However, it is not clear that low pulse rate processors would produce poorer speech recognition than high pulse rate processors. Original vocoder representations of speech (e.g., Hill et al., 1968) achieved high levels of speech recognition with only relatively slow temporal fluctuations below 20 Hz, which would imply that low pulse rates would be adequate. Shannon et al. (1995) showed no significant deterioration in speech recognition when speech envelope information was limited to below 50 Hz, which can be encoded by a pulsed carrier of little more than 100 pps.. And Stephan Brill (Brill, 1997) showed no difference in CIS speech processor performance in cochlear implants as he varied pulse rates from 200 to 4000 pps/electrode. Still, low-rate processors should be evaluated in actual ABI patients. If stimulation rates above 250 pps do yeild superior performance, then it may be possible to ~~e~~incorporate this into the design of the electrode array (e.g., use an array with many electrode sites, and cycle the stimul~~ate~~ation across those sites that appear to be at nearly the same point on the tonotopic axis of the VCN.)

We will implement CIS-type and SPEAK-type speech processors in present ABI patients with surface electrodes, and later in patients with the hybrid device, and measure speech recognition as a function of the pulse rate per electrode. Although we do not expect SIDNE to occur with surface electrodes, we should be able to evaluate whether or not speech processors with low pulse rates will produce unacceptable reductions in speech recognition. CIS and SPEAK-type processors will be implemented with 4 and 6 electrodes in ABI patients with present surface electrodes. Pulse rates of 50, 100, 200, and 400 pps per electrode will be used. With the penetrating electrodes, rates above 200 pps will be used only for a few minutes, to avoid possible SIDNE. Pitch estimates and will also be collected for these same pulse rates applied to individual electrodes at the medial and lateral locations on the surface electrode array. Modulation detection thresholds will be measured for 50, 100, 200, and 400 Hz modulation rates on 1000 pps carriers.

Speech Processor Fitting Protocol

Acoustic Simulation of ABI Speech Processor Fitting. We will perform a blinded study of speech processor optimization on normal-hearing listeners. One investigator will select four noise bands to mimic the tonotopic location and bandwidth of the ABI electrodes, similar to the conditions shown schematically in Figures 1 and 2. A noise-band processor implementing these bands will be programmed to run in real time on a Motorola 56003 DSP. A different investigator will then attempt to adjust the analysis bands on that subject to achieve the optimal performance. Techniques will be developed for first dividing the spectrum into equal log frequency portions, and then systematically altering the distribution of frequencies into each band to achieve the best speech recognition performance. If the "Mixed" condition is not significantly worse than the "Matched" condition this will simplify the procedure because only contiguous analysis bands will be considered. If the "Matched" condition is significantly better than the "Mixed" condition, then the additional parameter of the bandwidth of each filter must be added to the search space.

Processor Fitting Trials in Surface-electrode ABI Patients. In patients with the present surface electrode ABI there is a similar uncertainty about the absolute tonotopic location and bandwidth of the neural population stimulated by each electrode. In this device the electrodes are several mm distant from the stimulable neurons of the CN and so the spread of current is probably considerable. Nevertheless, the problems of matching the analysis filters to an unknown electrode configuration will provide experience with the same patient population and may help streamline the fitting protocol for eventual use on patients with penetrating electrodes. We will select present ABI patients who have a good range of pitch sensations across their electrodes. Analysis filters will be adjusted systematically in terms of their center frequency and bandwidth. The magnitudes of pitch estimates for the electrodes will be used as an initial estimate for dividing the speech spectrum, e.g., if only one electrode is high-pitch and the others are mostly low, a larger portion of the spectrum will be routed to the high-pitch electrode and the remaining spectrum will be divided among the low-pitch electrodes. We will try using all electrodes as well as using only electrodes that are significantly different in pitch.

Comparison of Surface and Penetrating electrode ABI systems

Pitch. To contrast the relative pitch salience of surface and penetrating electrode stimulation of the human cochlear nucleus we will measure and compare pitch magnitudes with the two electrode placements. When subjects are available with the implanted surface-penetrating hybrid system we will have them match the pitch of their surface electrodes against their penetrating electrodes and *vice versa*. The pulse rate on the standard electrode will be fixed at 50, 100, or 200 pps. The subject will be instructed to adjust the pulse rate on the comparison electrode to match the pitch of the standard. If no match can be achieved we will collect pitch estimates of the same pulse rates presented to the surface and penetrating electrodes.

Speech Recognition. When subjects are available with the implanted surface-penetrating hybrid ABI system, we will be able to directly compare speech processing strategies for surface electrodes with the same strategies on penetrating electrodes in the same individual. We will also measure speech recognition with processors strategies that use surface and penetrating electrodes together. Assignments of spectral regions to electrodes will be made on the basis of electrode pitch estimates. Because surface electrodes are expected to activate broader tonotopic regions, speech processors will be developed that use surface electrodes for low-acoustic frequency information, which can be conveyed primarily temporally, and use the penetrating electrode for higher-frequency spectral information that is conveyed primarily by tonotopic location. The frequency dividing these two spectral ranges will be parametrically adjusted to determine the best balance between spectral and temporal representations.

Psychophysical Methods

General Methods. Stimuli will be presented for detection or discrimination in a two-interval, forced-choice adaptive task. We will use a 3-down, 1-up adapting rule to converge on the stimulus value that produces 79.4% correct performance (Levitt, 1971).

In the first array of intranuclear microelectrodes, the lengths of microelectrodes will be staggered in intervals of 0.5 mm. This is based on studies of electrode interaction conducted in the cat (McCreery et al, 1998), the range of safe stimulus amplitude in the cats (McCreery et al, 1992, 1994, 1997) , and anatomical details of the human ventral cochlear nucleus. We will evaluate the accuracy of our estimate The Electrode interaction experiments will utilize forward masking, gap detection, and loudness balancing paradigms. Forward masking and gap detection will be measured in a two-interval, forced-choice adaptive task. Loudness balances will by measured using alternating loudness balance technique. This procedure presents a standard stimulus alternating with a comparison stimulus. The standard stimulus will consist of a single pulse train on a single electrode, while the comparison stimulus will consist of stimuli on two electrodes. The subject will be instructed to adjust the loudness of the standard stimulus until it's loudness is equal to that of the comparison stimulus. They will be instructed to bracket the loudness first by adjusting the standard stimulus to be clearly louder than, then softer than the comparison stimulus.

To understand the principle of using loudness summation to measure electrode interaction, consider the additivity of loudness of two acoustic noise bands. If the bands are widely spaced in frequency they stimulate two distinct regions of the auditory nerve. In this case the loudness of the two bands is the sum of the loudness values of each component band. If the two bands are within a critical band then the acoustic power of the two bands are summed and the loudness is determined by this increased in signal power. However, two bands will only increase the power by 3 dB, resulting in only a small increase in loudness. Thus, acoustic loudness summation ranges between a small loudness increase (complete overlap of spectral range) to double the loudness (no overlap). Partial overlap of the spectral region of the two bands (and thus of the neural populations excited) will result in intermediate values of loudness summation. The loudness summation in electrical stimulation is similar to that in the acoustic case. If the two electrodes stimulate non-overlapping neural populations, the loudness of the two-electrode stimulus is the sum of the perceived loudness of each electrode by itself. If the electrodes stimulate the same neural population then the nerves are stimulated by the sum of the electrical waveforms from the two electrodes. For example, if each electrode is presenting a 100 Hz biphsic pulse train and they are interleaved in time, then the neurons are stimulated by a 200 Hz pulse train. And the loudness of a 200 Hz pulse train is the same as the loudness of a 100 Hz pulse train. Thus, compared to the loudness of a single electrode stimulus, the increase in loudness of a two-electrode stimulus will range between no increase (complete interaction) and double the loudness (no interaction). Partial overlap of the neural populations excited by the two electrodes will result in intermediate values of loudness summation.

<u>Acoustic Simulations of Implant Speech Processors</u>**.** The signals are digitized at a 10 kHz sample rate and passed through a pre-emphasis filter to whiten the spectrum (6 dB/octave decrease below 1200 Hz). The signal is then split into frequency bands (6th order Butterworth filters). The envelope is extracted by half-wave rectification and low-pass filtering (6 dB/octave filters with cut-off frequencies of 50, 160 or 500 Hz). The envelope from each band is then used to modulate a wide-band white noise. The modulated noise is frequency-limited by filtering. In some conditions the filters will be the same filters used in the analysis bands, while in other conditions the band-pass filters will be different in center frequency and bandwidth from the analysis filters. The band-pass

filtering reduces the modulation depth to some degree because it removes the modulation side-bands. The resulting modulated noises from each band are combined, low-pass filtered at 4 kHz, amplified (Crown D75) and presented to the listener through headphones (TDH-49).

Speech Materials. Phoneme recognition performance will be assessed using two measures: medial vowels and consonants. Vowel recognition is measured in a 12-alternative identification paradigm, including 10 monophthongs and 2 diphthongs, presented in a /h/-vowel-/d/ context (heed: h/i/d, hawed: h/ /d, head: h/E/d, who'd: h/u/d, hid: h/I/d, hood: h/U/d, hud: h/ /d, had: h/Q/d, heard: h/_/d, hoed: h/o/d, hod: h/A/d, hayed: h/e/d). The tokens for these closed-set tests were digitized natural productions from 5 men, 5 women, 3 boys, and 2 girls, drawn from the material collected by Hillenbrand et al. (1995). Consonant recognition is measured in a 16-alternative identification paradigm, for the consonants /b d g p t k l m n f s v z j /, presented in a /a/-consonant-/a/ context. Two repetitions of each of the 16 consonants were produced by three speakers (1 male, 2 female) for a total of 96 tokens (16 consonants * 3 talkers * 2 repeats).

A stimulus token is randomly chosen from all 180 tokens in vowel recognition and from 96 tokens in consonant recognition and presented in random order. Following the presentation of each token, the subject responds by pressing one of 12 buttons in the vowel test and one of 16 buttons in the consonant test, each marked with one of the possible responses. No feedback is provided, and subjects are instructed to guess if they're not sure, although they're cautioned not to provide the same response for each guess. The subjects start the formal test without training and with no appreciable period of adjustment to the new processor. Data will be collected from two runs by each subject, which represents 30 presentations of each vowel and 12 presentations of each consonant.

Recognition of words in sentences will be measured using sentences from the ~~h~~Hearing in Noise test (HINT: Nilsson et al., 1994).

SUMMARY

Measurement of basic psychophysical abilities and speech recognition performance of patients with penetrating-electrode ABIs must be compared to the same measures with surface-electrode ABIs. To offset the increased risks of penetrating electrodes over surface electrodes, performance must be demonstrably better with penetrating electrode systems. The present progress report summarizes issues in psychophysical capabilities and in speech processor design that will allow direct comparison of surface and penetrating electrode ABIs.

REFERENCES

Boex, C.S., Eddington, D.E., Noel, V.A., Rabinowitz, W.M., Tierney, J., and Whearty, M.E. (1997). Restoration of normal loudness growth for CIS sound coding strategies, **Abstracts of the 1997 Conference on Implantable Auditory Prostheses**, Asilomar, CA, August, page 26.

Brackmann, D.E., Hitselberger, W.E., Nelson, R.A., Moore, J.K., Waring, M., Portillo, F., Shannon, R.V., and Telischi, F. (1993). Auditory Brainstem Implant. I: Issues in Surgical Implantation, **Otolaryngology, Head and Neck Surgery, 108, 624-634.**

Brill, S. (1996). Optimization of channel number and stimulation rate in the COMBI 40+, **International Workshop on Cochlear Implants**, Vienna, Austria, October 24-25.

Chatterjee, M. and Shannon, R.V. (1998). Forward masking excitation patternsas a measure of electrode interaction in cochlear implants: dependence on stimulus parameters. **Journal of the Acoustical Society of America**, 103(5), in press.

Cohen, L.T., Busby, P.A., Whitford, L.A. and Clark, G.M. (1996a). Cochlear implant place psychophysics: 1. Pitch estimation with deeply inserted electrodes, **Audiology & Neuro-Otology**, 1, 265-277.

Cohen, L.T., Busby, P.A., and Clark, G.M. (1996b). Cochlear implant place psychophysics: 2. Comparison of forward masking and pitch estimation data, **Audiology & Neuro-Otology**, 1, 278-292.

Collins, L.M., Zwolan, T.A., and Wakefield, G.H. (1998). Comparison of electrode discrimination, pitch ranking, and pitch scaling in postlingually deafened adult cochlear implant subjects, **J. Acoust. Soc. Amer.**, in press.

Daniloff, R.G., Shiner, T.H., and Zemlin, W.R. (1968). Intelligibility of vowels altered in duration and frequency, **J. Acoust. Soc. Amer.**, 44, 700-707.

Delhorne, L.A., Eddington, D.E., Noel, V.A., Rabinowitz, W.M., Tierney, J., and Whearty, M.E. (1997). Speech reception with CIS processing: Longitudinal evaluations and implications from acoustic simulations, **Abstracts of the 1997 Conference on Implantable Auditory Prostheses**, Asilomar, CA, August, page 49.

Dorman, M.F., Loizou, P.C. and Rainey, D. (1997a). Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs, **J. Acoust. Soc. Amer.**, 102, 2403-2411.

Dorman, M.F., Loizou, P.C. and Rainey, D. (1997b). "Simulating the effect of cochlear-implant electrode insertion-depth on speech understanding", **J. Acoust. Soc. Amer.**, 102(5), 2993-2996.

Fishman, K., Shannon, R.V., and Slattery, W.H. (1997). Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor, **J. Speech Hear. Res.**, 40, 1201-1215.

Fu, Q.-J., Shannon, R. V., Zeng, F.-G., & Chatterjee, M. (1996). Electrode interactions measured by loudness summation in cochlear implant listeners, **Journal of the Acoustical Society of America,** 100, 2631.

Fu, Q.-J. and Shannon, R.V. (1998a). Effects of electrode location and spacing on speech recognition with the Nucleus-22 cochlear implant, **Journal of the Acoustical Society of America,** 104, (in press).

Fu, Q.-J. and Shannon, R.V. (1998b). Effects of amplitude nonlinearities on speech recognition by cochlear implant users and normal-hearing listeners, **Journal of the Acoustical Society of America**, 104, (in press).

Fu, Q.-J., Shannon, R.V., and Wang, X. (1998). The effect of noise on speech recognition by normal-hearing and cochlear implant listeners, **Journal of the Acoustical Society of America**, (submitted October 29, 1997).

Hillenbrand J., Getty, L.A., Clark, M.J., and Wheeler K. (1995). Acoustic characteristics of American English vowels, **J. Acoust. Soc. Am.** 97, 3099-3111.

Hanekom J.J., and Shannon R.V. (1996). "Place pitch discrimination and speech recognition in cochlear implant users", **South African Journal of Communication Disorders,** 43, 27-40.

Gap detection as a measure of electrode interaction in cochlear implants: implications for speech processor fitting. Submitted to **Journal of the Acoustical Society of America**, June 1997.

Hill, F.J., McRae, L.P., and McClellan, R.P. (1968). Speech recognition as a function of channel capacity in a discrete set of channels, **J. Acoust. Soc. Amer.**, 44, 13-18.

Levitt, H. (1971) Transformed up-down methods in psychoacoustics, **J. Acoust. Soc. Amer**., 49, 467-477.

Lim, H.H., Tong, Y.C., and Clark, G.M. (1989). Forward masking patterns produced by intracochlear stimulation of one and two electrode pairs in the human cochlea, **J. Acoust. Soc. Am.**, 86, 971-980.

Masterton, R.B. (1992). Role of the central auditory system in hearing: the new direction. **Trends in Neuroscience**, 15(8), 280-285.

McCreery, D.G., Yuen, T.G.H., Agnew, W.F., and Bullara, L.A. (1997). A characterization of the effects on neuronal excitability resulting from prolonged microstimulation with chornically implanted microelectrodes, **IEEE Trans. Biomed. Eng.**, 44, 831-939.

McCreery, D.B., Yuen, T.G.H., Agnew, W.F., and Bullara, L.A. (1992) Stimulation with chronically implanted microelectrodes in the cochlear nucleus of the cat: Histologic and physiologic effects. **Hearing Res.** 62:42-56.

McCreery, D.B., Yuen, T.G.H., Agnew, W.F. and Bullara, L.A. (1994) Stimulation parameters affecting tissue injury during microstimulation in the cochlear nucleus of the cat **Hearing Res.** 77:105-115 .

McCreery, D.B. Shannon R.V., Moore J.K, Chattegee, M. (1998) Accessing the tonotopic organization of the ventral cochlear nucleus by intranuclear microstimulation (to be publised in **IEEE Trans. Reheb .Eng. ,** December 1998)

Nagafuchi, M. (1976). Intelligibility of distorted speech sounds shifted in frequency and time in normal children, **Audiology**, 15, 326-337.

Nilsson, M.J., Soli, S.D., and Sullivan, J. (1994). Development of a hearing in noise test for the measurement of speech reception threshold, **J. Acoust. Soc. Am**., 95, 1085-1099.

Otto, S.A., Shannon, R.V., Brackmann, D.E., Hitselberger, W.E., Staller, S., and Menapace, C. (1998). The multichannel auditory brainstem implant: Performance in 20 patients, **Otolaryngology, Head and Neck Surgery**, 118(3), 291-303.

Rosen, S. (1992). Temporal information in speech and its relevance for cochlear implants, in **Philos. Trans. Royal Soc. London Ser. B Biol. Sci.**, 336, 367.

Shannon, R.V. (1983). Multichannel electrical stimulation of the auditory nerve in man: II Channel interaction. **Hearing Research**, 12, 1-16.
Shannon, R.V. (1985). Loudness summation as a measure of channel interaction in a multichannel cochlear implant. In Cochlear Implants, R.A. Schindler and M.M. Merzenich, Eds., Raven Press, New York.

Shannon, R.V., Zeng, F-G., and Wygonski, J. (1992). Speech recognition using only temporal cues, in **The Auditory Processing of Speech: From Sounds to Words**, M.E.H. Schouten (Ed.), Mouton-DeGruyter, New York, pp 263-274.

Shannon, R.V., Fayad, J., Moore, J.K., Lo, W., O'Leary, M., Otto, S., and Nelson, R.A. (1993). Auditory brainstem implant. II: Post-surgical issues and performance, **Otolaryngology, Head and Neck Surgery,** 108, 635-643.

Shannon, R.V., Zeng, F.-G., Wygonski, J., Kamath, V., and Ekelid, M. (1995). Speech recognition with primarily temporal cues, **Science,** 270, 303-304.

Shannon, R.V., Moore, J., McCreery, D., and Portillo, F. (1997). Threshold-distance measures from electrical stimulation of human brainstem**, IEEE Trans. on Rehabilitation Engineering,** 5, 1-5.

Shannon, R.V., Zeng, F.-G., and Wygonski, J. (1998). Speech recognition with altered spectral distribution of envelope cues, **J. Acoust. Soc. Amer.**, 104, in press.

Tiffany, W.R., and Bennett, D.A. (1961). Intelligibility of slow-played speech, **J. Speech Hearing Res.**, 4, 248-258.

Van Tasell, D.J., Soli, S.D., Kirby, V.M., and Widin, G.P. (1987) Speech waveform envelope cues for consonant recognition, **J. Acoust. Soc. Amer.**, 82, 1152-1161.

Zwolan, T.A., Collins, L.M., and Wakefield, G.H. (1997). Electrode discrimination and speech recognition in postlingually deafened adult cochlear implant subjects, **J. Acoust. Soc. Amer.**, 102(6), 3673-3685.